University of Reading

Explaining the diversity in Malay-English code-switching patterns: the contribution of typological similarity and bilingual optimization strategies

Article

Updated Version

Creative Commons: Attribution 4.0 (CC-BY)

Open Access

Treffers-Daller, J. ORCID: https://orcid.org/0000-0002-6575-6736, Majid, S., Yap, N. T. and Flynn, N. ORCID: https://orcid.org/0000-0002-3710-9910 (2022) Explaining the diversity in Malay-English code-switching patterns: the contribution of typological similarity and bilingual optimization strategies. Languages, 7 (4). 299. ISSN 2226-471X doi: 10.3390/languages7040299 Available at https://centaur.reading.ac.uk/109680/

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.3390/languages7040299

Publisher: MDPI

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in

the End User Agreement.

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

languages

MDPI

Article

Explaining the Diversity in Malay-English Code-Switching Patterns: The Contribution of Typological Similarity and Bilingual Optimization Strategies

Jeanine Treffers-Daller [1,*], Sheikha Majid [2,*], Yap Ngee Thai [3] and Naomi Flynn [4]

[1] Department of English Language and Applied Linguistics, University of Reading, Reading RG6 6EL, UK
[2] Faculty of Education, Universiti Teknologi MARA, Puncak Alam 42300, Malaysia
[3] Department of English, Universiti Putra Malaysia, Serdang 43400, Malaysia
[4] Institute of Education, University of Reading, Reading RG1 5EX, UK
* Correspondence: j.c.treffers-daller@reading.ac.uk (J.T.-D.); sheikha@uitm.edu.my (S.M.)

Abstract: Bilingual speakers often engage in code-switching, that is the use of lexical items and grammatical features from two languages in one sentence. Malaysia is a particularly interesting context for the study of code-switching because Malay-English code-switching is widely practiced across formal and informal situations, and the available literature reveals that there is a great diversity in switch patterns in this language pair. One of the most remarkable characteristics of Malay-English code-switching is the high frequency of switches of function words (pronouns, modal verbs, demonstratives, etc.), which is very unusual in most code-switching corpora. Here, we analyse the structural properties of Malay-English code-switching, which have received less attention than functional analyses in the academic literature on code-switching in this language pair. We first summarize the literature on the different types of code-switching that are found in a range of sources, and then analyze the code-switching patterns in the speech of two teachers of English in Malaysia. We conclude with a discussion of the variables that can explain the diversity found, in particular structural factors (similarity between the word orders of both languages, and the limited number of inflections), and bilingual optimization strategies, as well as strategies of neutrality and efficiency.

Keywords: code-switching; codemixing; English; Malay; bilingual; congruent lexicalization; function words

check for updates

Citation: Treffers-Daller, Jeanine, Sheikha Majid, Yap Ngee Thai, and Naomi Flynn. 2022. Explaining the Diversity in Malay-English Code-Switching Patterns: The Contribution of Typological Similarity and Bilingual Optimization Strategies. *Languages* 7: 299. https://doi.org/10.3390/languages7040299

Academic Editors: Raquel Fernández Fuertes and Juana M. Liceras

Received: 6 July 2022
Accepted: 15 November 2022
Published: 23 November 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

In most bilingual communities, speakers switch between two languages whenever the situation requires or allows it. This behaviour, which is generally referred to as code-switching, is also very common in Malaysia (David 2003; David et al. 2009a; Ozog 1987). As code-switching has been studied extensively since the earliest publications in the second half of the twentieth century (N. Abdullah 1975; Clyne 1967, 1987; Lehtinen 1966; Pfaff 1979; Poplack 1980, it is now commonplace that bilinguals can switch in many different ways depending on (a) typological differences between languages; (b) sociolinguistic factors such as the relative status of different languages, the setting in which a conversation takes place and the topic that is being discussed; (c) personal characteristics of the interlocutors (e.g., relative status) in a conversation, and (d) psycholinguistic variables, including informants' ability to inhibit non-target languages and to monitor language choice during speech production. However, it is clear that many more variables impact on switch patterns, and that some are the result of the conventionalisation of basically arbitrary patterns (Muysken 2000).

We have chosen Malay-English as the language pair to be studied because, first of all, there is evidence from earlier studies that code-switching is widely practised in formal and informal settings in Malaysia (N. Abdullah 1975; Jacobson 2001; Ozog 1987). It has

been studied frequently in educational contexts (see Majid 2019), where the medium of instruction is officially English, but in practice staff and students often rely on code-switching to facilitate understanding of school language, as is the case in many contexts where English Medium Instruction has been introduced in schools (Dearden 2014). Second, several of these earlier studies (e.g., Ozog 1987) show there are many instances of switches of function words in Malay-English code-switching, which is highly unusual in code-switching in most language pairs (Muysken 2000). Finally, it is important for new evidence to be shared with the research community from multilingual contexts outside Europe and North-America, which continue to dominate the field. In Western societies code-switching is a stigmatised form of language behaviour (Badiola et al. 2018; Dewaele and Li 2014; Koban 2016; Jaworska and Themistocleous 2018; Poplack 1980). In non-Western societies, by contrast, there seems to be less of a stigma attached to this form of language behaviour (Auer et al. 2014). This is also the case in Malaysia, where code-switching is the norm among Malay-English bilinguals belonging to different ethnic groups (David et al. 2009a). It is possible that in communities where code-switching is the norm, more original (or perhaps intimate) forms of code-switching appear and spread through the community than in contexts where it is highly stigmatised. Indeed, the available evidence suggests that there is a wide range of different types of code-switching in Malaysia. Particularly intriguing is the fact that switching of function words seems to be much less restricted in this language pair than in other language pairs. The diversity of the patterns and the frequency of switches of function words call for an explanation. A key aim of the current paper is therefore to provide a detailed overview of the structural properties of Malay-English code-switching as found across a range of formal and informal contexts, and to offer explanations for the diversity of the patterns. In addition, we analyse a new data set of code-switching in an educational context, collected by Majid (2019), and compare the findings against those from the available literature.

We start by giving an overview of Muysken's model (Section 2.1), and the ways in which different types of code-switching can be distinguished (Section 2.2). After this we summarize the available literature on code-switching in Malaysia (Section 2.3). Then, follows a section on the current project, with aims, research questions, and methods (Section 3). After this, we describe and discuss the findings (Section 4) and we finish with a conclusion (Section 5).

2. Different Types of Code-Switching

2.1. Muysken's (2013) Four-Way Code-Switching Typology

For the purposes of the current paper we base our analyses on Muysken's (2013) four-way typology of intrasentential code-switching, as this is based on detailed analyses of over forty different contact situations, and considers both linguistic, sociolinguistic and psycholinguistic variables. It is therefore more clearly rooted in empirical evidence than any other model. One of the most common forms of code-switching is INSERTION (INS), which involves the embedding of a content word or a phrase from language A into a stretch of speech from language B, as in (1), where the Turkish noun *maaş* 'salary' is embedded into a German prepositional phrase, and is allocated German gender and case marking on the determiner.

(1)	Bist	du	mit	dem	*maaş*	zufrieden?
	Are	you	with	the.masc.DAT	salary	content

'Are you content with the salary?' (Treffers-Daller 2020, p. 244)

The type of code-switching in (1) is what Myers-Scotton (1998) calls classic code-switching, in that the division of labour between the languages is asymmetrical: there is a clear matrix language (ML), which provides the grammatical frame. In (2) the ML is likely to be German, while Turkish is the embedded language (EL)—but see below for further discussion.

The second type of code-switching consists of switches of longer stretches of speech (not just individual content words or short phrases). When a longer stretch of speech in language A alternates with a longer stretch of speech in language B, and there is only a loose relationship between the parts in each language, this is called ALTERNATION (ALT), as in (2) where the switch takes place between two co-ordinated clauses.

(2)

Nadine est	née	au	mois	d'avril	e*n*	*dan*	*in de maand*
Nadine was	born	inthe	month	of April	and	then	in the month
Oktober heb		*ik*	*een winkel*	*open+gedaan*			*in* . . .
October have		I	a shop	open.PTCP. do.PTCP			in..

'Nadine was born in April, and then I opened a shop in October.' (Treffers-Daller 1994, p. 30).

The third form of code-switching is BACKFLAGGING (BFL), which was treated as a subtype of ALT in Muysken (2000), where a tripartite division was proposed rather than the four-way typology proposed in Muysken (2013). This involves the attachment of a discourse marker or co-ordinate conjunction from a bilingual's first language to a stretch of speech in the speaker's second language, as in (3), where Malay particle *lah*, is attached to an English utterance (see also a discussion of this particle in Section 2.3.1). In all Malay-English examples, English is given in italics and Malay in regular font. We have kept the authors' own glosses in all cases.

(3) *Buy this* lah
Buy this DM
'Buy this.' (Tay et al. 2016, p. 490)

The fourth type of code-switching is CONGRUENT LEXICALIZATION (CLX). This type of code-switching can be found in utterances where the grammar and the lexis of both languages interact. This may involve switching of function words as well as content words, and these are combined in a stretch of speech for which the word order is often shared between both languages, as in the German-English example (4), where the English expression *make friends with* is translated into German and partially filled with English words. Mixed collocations such as this are typical of CLX (Muysken 2000).

(4)

Wir hab+en	*friends*	ge+mach+t	mit dem	*shopowner*
We have+PL	friends	PTCP+make+PTCP	with the-DAT.SG	shopowner

'We have made friends with the shopowner.' (Hofweber et al. 2016, p. 651).

The other reason why this qualifies as CLX is because the speaker uses English word order in the second half of the utterance: the prepositional phrase (PP) *mit dem shopowner* 'with the shopowner' appears to the right-hand side of the past participle *gemacht* 'made'. While the postverbal position is a canonical order for PPs in English, in German PPs preferably occur before the verb. Postverbal PPs are less common, and typically occur under specific pragmatic conditions (Averintseva-Klisch 2009). However, it is a possible word order in both languages. Thus, word order is shared between both languages at this point.

This type of code-switching has received very little attention in the literature, and its very existence is doubted by many researchers. We hope to provide further evidence that congruent lexicalization (CLX) occurs frequently in the Malay-English code-switching data set under study. While this type of code-switching is not expected among languages which are typologically quite distinct, its occurrence may have been facilitated by a variety of factors, including the structural properties of both languages and the depth of contact between Malay and English.

CLX resembles Myers-Scotton's (1998) composite ML but differs from it because Myers-Scotton (1998, p. 292) defines composite ML as follows: "the ML is a composite of lexical structure from two or more sources". Thus, only *lexical* items from both languages appear in clause with a composite ML, while *functional* items can only come from one

language in the clause. By contrast, CLX may involve the interaction of functional as well as lexical items from both languages within one clause.

In their review of Muysken (2000), Poplack and Walker (2003) note that Muysken's typology is novel because the field of code-switching had hitherto mainly distinguished between two approaches which could be characterized as insertional and alternational. The key issue in studies focused on insertional patterns is the *switched element* and its characteristics, while those who work on alternational patterns are mostly interested in *switch points*, that is the transition point between two languages in an utterance. Myers-Scotton (1993, et seq) exemplifies the insertional approach, as in her work, a fundamental distinction is made between an EL (A) and a ML (B), and the aim is to identify which elements from language A can be inserted into the grammatical frame of language B. While there are different views on how the ML should be determined, one widely used approach is that the inflection on the verb determines the ML because it is one of the highest functional heads in the syntactic tree (see Deuchar 2020 for fuller discussion). However, as Malay has very few inflections[1] (Prentice 1990), we cannot use the criterion of inflection on the verb to determine the ML in Malay-English code-switching, if the verb is Malay. Instead, we use another criterion, namely the language of the base form of the verb. This criterion was chosen because verbs are the 'semantic kernel' of the sentence (Muysken 2000, p. 67) in that they assign semantic roles in the clause.

While in some forms of code-switching distinguishing between ML and EL is very useful, this is not always the case. In switches which consist of alternation between longer stretches of speech belonging to either language A or language B, as in the work of Poplack and colleagues (Poplack 1980, et seq), there is no embedding of material of the EL into an ML. As Deuchar et al. (2007) explain, a key contribution of Muysken (2000) is to bring these two different approaches together in one typology, thus highlighting the diversity of switching patterns in different communities.

Another important distinction between Poplack's and Myers-Scotton's approaches is that Poplack does not consider single words from language A in language B as code-switches but as borrowings (Poplack 2018). Thus, *maaş* 'salary' in (1) does not qualify as a code-switch under this approach. By contrast, switches which consist of more than one word, such as the switch in (2), where the switch takes place between two coordinated clauses is a genuine code-switch in Poplack's approach. In Myers Scotton's work, however, the distinction between code-switching and borrowing is not seen as fundamental, and single content words are potential code-switches. We take the view that the four strategies distinguished by Muysken underlie *both* code-switching and borrowing, and that in a paper which aims at studying these four strategies, distinguishing between the two is not pertinent to the argumentation. Analysing the morphosyntactic integration of English words in Malay (and vice versa) to establish whether a word is borrowed is also very difficult, first of all because both languages have the same basic word order (SVO). This means the surface word order is very similar in most structures, except for the NP, because adjectives appear after the noun in Malay and before the noun in English. Second, there is very little inflection in Malay. Thus, almost all single word switches/borrowings appear in their bare form. For these two reasons, studying the morphosyntactic integration of the phenomena would also require obtaining monolingual stretches of speech from the informants in both languages against which the bilingual data can be compared. However, in a context where code-switching is a community-wide discourse mode, it can be difficult to obtain such monolingual samples. In the absence of such evidence, it is not currently possible to engage with the issue of the distinction between borrowing and code-switching in our data. The reader is referred to Deuchar (2020) and Treffers-Daller (2005, Treffers-Daller) for a detailed review. For the purposes of the current paper, we follow Muysken (2014, p. 254) who proposes borrowing results from insertion, but the two should not be equated.

We will now illustrate how the diagnostic features can be used to differentiate CLX from other types of code-switching using examples from the available literature.

2.2. Differentiating between CLX, INS and ALT

Importantly, the different types of intrasentential code-switching differ only *gradually* from each other, as they share some feature settings too. As can be seen in Table 1, the features 'single constituent' and 'morphological integration' are characteristic for INS, and marked with a plus sign (+). For ALT, 'single constituent' is neutral (because alternations may or may not consist of one constituent) and therefore marked with 0, and 'morphological integration' is counter-indicative for ALT, and therefore marked with a minus sign (−).

Table 1. Diagnostic features of the three patterns of code-mixing (Muysken 2000, p. 230), augmented with the category of backflagging (Muysken 2013).

Diagnostic Feature	Insertion	Alternation	Congruent Lexicalization	Backflagging
CONSTITUENCY				
Single constituent	+	0	0	+
Several constituents	−	+	0	−
Non-constituent	−	−	+	−
Nested aba	+	−	0	−
Non-nested aba	−	+	+	+
ELEMENT SWITCHED				
Diverse switches	−	0	+	−
Long constituent	−	+	−	−
Complex constituent	−	+	−	−
Content word	+	−	−	−
Function word	−	−	+	−
Adverb, conjunction	−	+	−	+
Selected element	+	−	+	−
Emblematic or tag	−	+	0	+
SWITCH SITE				
Major clause boundary	0	+	0	+
peripheral	0	+	0	+
Embedding in discourse	0	+	0	+
Flagging	−	+	−	0
Dummy word insertion	+	0	−	−
Bidirectional switching	−	+	+	+
PROPERTIES				
Linear equivalence	0	+	+	+
Telegraphic mixing	+	−	−	−
Morphological integration	+	−	+	−
Doubling	−	+	−	−
Triggering	0	0	+	−
Mixed collocations	0	−	+	−
Self-corrections	−	+	−	0

Reproduced from Cacoullos, Dion, Lapierre_Linguistic Variation, 1st Edition by Rena, Torres Cacoullos; Nathalie, Dion; André, Lapierre, published by Routledge. ©Routledge, 2015, reproduced by arrangement with Taylor & Francis Group.

CLX is different from the other three types of code-switching on a number of features. First of all, it may involve switching of elements that are not necessarily constituents (NON-CONSTITUENT SWITCHING or RAGGED SWITCHING), as in (5), where we find a mixture of French and Dutch expressions. The French expression *avoir quelque chose en horreur* and its Dutch equivalent *afschuw hebben van* both mean 'have a horror of, to loathe sth.' In (5) these two expressions are combined in a mixed collocation because the speaker starts the expression in Dutch with *had* 'had' (the past tense form of *hebben* 'to have') and uses the Dutch determiner *de* 'the' before switching to French for the compound noun *chou rave* 'kohlrabi'. This is an example of non-constituent switching, because the switch does *not* take place at a major constituent boundary between the verb and the NP, but at a later point within the NP. The chunk *chou rave en horreur* is not a full constituent either, but an intermediate projection between the noun and the full NP. Finally, the use of the Dutch article is noteworthy, because as Muysken (2000, p. 103) points out, in Dutch there are no articles in generic plurals; in French, however, it is common to use plural forms for generic uses of nouns. In other words, in (5), the Dutch article is used in a French way. This very intimate form of mixing two languages is probably best analysed as CLX.

(5) Ik had de *chou rave* *en horreur*
I had the kohlrabi in horror
'I hated kohlrabi.' (Treffers-Daller 1994 analysis in Muysken 2000, p. 103).

Second, it is typical for switching of the CLX type to be rather diverse, in that switches belonging to different syntactic classes (not just nouns or other lexical material) are found in a data set. Third, CLX can involve switching of function words, which is generally quite restricted in code-switching (Lehtinen 1966), because of the lack of categorical equivalence between function words from different languages (Muysken 2011). An example from Sranan-Dutch is given in (6), where the Dutch expression *draad oppakken*, 'take over' is given in italics and the Sranan function words in regular font. Note that this is again a mixed collocation because of the presence of a Sranan determiner in the middle of the Dutch expression. The other function words (a pronoun and a future marker) are also in Sranan.

(6) mi o *pak* a *draad* *op*
I FUT pick the thread up
'I will take over.' (Bolle 1994; in Muysken 2000, p. 141).

Third, TRIGGERING (Clyne 1967), and fourth, switches which consist of mixed collocations, as in (4) and (5) are typical for CLX. Triggering can happen when words that are the same or very similar in both language, such as the cognates in (4) are activated. Such trigger words may activate other words from the same language, which can, in turn, lead speakers to continue in a language different from the one in which they had started the utterance, as in (7) from Clyne (2003), in which a German-English bilingual who starts her utterance in German, switches to English for the title of the program *the Nelsons*, which triggers a switch to English for the remainder of the utterance.

(7) Am Montag seh' ich am liebsten, *the Nelsons' and then doctor Kildare and then we*
On Monday watch I preferably the Nelsons and then doctor Kildare and then we
turn it off.
turn it off.
'Mondays, I like watching the Nelsons and then doctor Kildare and then we turn it off.'
(Clyne 2003, p. 75, informant talking about preferred television programs)

As shown in Table 1, there are also similarities between the four types of code-switching. On the one hand, INS and CLX share the feature that switches may consist of selected elements (that is complements of verbs or prepositions), as in (5) where the PP *en horreur* 'in horror' is selected by the Dutch verb *had* 'had', and both types of code-switching can involve morphological integration (if the morphological properties of the languages involved allow for it). On the other hand, CLX is different from INS, as for INS

the switched element is nested a b a, which means that the elements before and after the switch are grammatically related, as in (4) where *friends* is surrounded by two German words: the auxiliary *haben* 'have' and the past participle *gemacht* 'made', which are clearly grammatically related[2]. Thus, *friends* satisfies the criterion of nestedness and is likely to be an insertion. Such nestedness is not typical for CLX.

The opposite (non-nested a b a) can be seen in (8), which starts in Dutch with the PP *bij mijn broer* 'at my brother's', after which the speaker switches to French for the core part of the sentence, and finishes with the tag *en alles* 'and everything', Importantly, although the French part is surrounded by Dutch expressions, there is no grammatical link between the PP *bij mijn broer* 'at my brother's' before the switch and the tag *en alles* after the switch.

(8) *Bij mijn broer*, y a un ascenseur *en alles*
At my brother's there is a lift and everything
'At my brother's house, there is a lift and everything.' (Treffers-Daller 1994, p. 221)

Thus, the switches in (8) are unlikely to be INS but the utterance could be seen as containing two alternations (the PP and the tag).

What is missing from the typology so far is a weighting of the different features. As Muysken (2014, p. 248) points out, some features could be more important than others. Utterances where different features appear to conflict with each other (e.g., with some features favouring INS and others ALT), could help to develop a hierarchy of importance of the diagnostic features. In (9), for example, the Dutch PP *op mijn gemakske* 'at my ease', could be interpreted as INS, because there is a grammatical relation between the French verb *étais* 'was' which precedes the switch and the French construction *en train de regarder les étoiles* 'in the process of watching the stars', which follows it. In other words, the Dutch PP is nested a b a in a French construction, and the feature 'nestedness', which is indicative of INS, should be set to positive. However, we may reach a different conclusion if we note that this PP is a manner adverbial. This means the PP is an adjunct rather than a complement, and the feature 'selected element' should be set to negative, which points in the direction of ALT.

(9) J'étais au balcon *op mijn gemaks+ke* en train de regarder les étoiles
I was on the balcony at my ease+DIM in process of watching the stars
'I was on the balcony at my ease watching the stars.' (Treffers-Daller 1994, p. 131)

In the case of (9), we would prefer to see this switch as ALT, because switches of adverbs are very frequent in the data set, which makes ALT the dominant pattern in the French-Dutch data from Brussels (Muysken 2000). Further research will need to indicate whether selection is a stronger criterion for INS than nestedness.

Because the features characterizing CLX overlap partly with those of other code-switching strategies one might wonder if CLX does indeed constitute a separate code-switching type. In fact, Muysken (2014) suggests that the key distinction is between INS and ALT and points to the centrality of language distance in the discussion about the typology of code-switching, saying:

> 'Congruent lexicalization is the epiphenomenal result of code-switching under the specific circumstances of similarity between the languages involved rather than a strategy in its own right.' (Muysken 2014, p. 256)

There is considerable evidence that CLX does indeed occur when languages are similar to each other either in grammar or vocabulary. In (4), there are five German/English cognates: *wir/we*; *haben/have*; *friends/Freunde*; *gemacht/made*; and *dem/the*. In the context where (4) was collected (among heritage speakers of German in South Africa) there is also a long tradition of language contact, as a result of which convergence between the two language systems may have taken place, and this is likely to be an additional key condition for CLX to arise. In fact, the duration and intensity of contact appears to be more important than typological similarity, because in a corpus of Sranan-Dutch code-switching

there are many examples of CLX (Bolle 1994; reported in Muysken 2000), even though the two languages belong to different language families.

If indeed similarity is a key condition for CLX to take place, this raises another issue, namely whether similarity between languages is an objective fact, to be operationalised by measuring the Levenshtein distance (Levenshtein 1966) between two languages, for example, or whether the distance between languages is in the eye of the beholder, and *perceptions* of the relatedness of languages play a more important role in determining the distance between languages (see also the discussion about psychotypology in Second Language Acquisition (e.g., Kellerman 1983). The fact that CLX is found even in languages which are—objectively—typologically clearly distinct seems to suggest the latter is the case. That similarity between languages facilitates code-switching has been discussed at an earlier stage in the code-switching literature, for example by Clyne (1987) under the heading CONGRUENCE. As Sebba (1998, p. 7), puts it: 'the locus of congruence is in the mind of the speaker', and bilinguals 'create' congruent categories by finding common ground between the languages concerned. An example of the creative construction of congruence can be found in the utterance from a Turkish-German bilingual who switches in the middle of a relative clause despite the total lack of 'objective' similarity between the two languages in formation of relative clauses (see Treffers-Daller 2020 for details).

Muysken's (2013) typology is innovative for a variety of reasons. First and foremost, it shows that there is a form of code-switching (CLX) for which the two languages are not kept strictly apart. This is new by comparison with widely held beliefs that code-switching involves *a complete shift* from one language to another for a word, a phrase or a whole sentence (Grosjean 2001; McClure 1977; Poplack and Meechan 1995). As in CLX the grammars and the lexicons of both languages can interact, the boundaries between the languages are blurred, and in many cases no ML can be distinguished (but see Deuchar et al. 2007 for solutions to the issue of determining the ML). According to Muysken (2000, 2014) CLX may be seen as akin to LANGUAGE VARIATION and STYLESHIFTING, that is changes in pronunciation, lexical choice or grammar depending on the formality of the social setting (Labov 1972), but it is also possible to see CLX as being similar to CROSS-LINGUISTIC INFLUENCE (Smith and Kellerman 1986). This can clearly be seen in (4) and (5), where there is not only importation of lexical material from one language into another, but also cross-linguistic influence at the level of the grammar. However, crosslinguistic influence cannot be equated with code-switching as, as for the former there is generally no lexical material being transferred from one language to another, while that is necessarily the case in code-switching (see also Treffers-Daller 2009 for fuller discussion).

The second reason why Muysken's typology is innovative is because it is explicitly linked to sociolinguistic and psycholinguistic variables. As noted above, CLX is expected in contexts where there is a long tradition of language contact, and there are structural parallels between languages (e.g., closely related languages). Over time, such parallels may have increased through convergence. Thus, what is ungrammatical in the standard varieties of English (e.g., omission of articles) may be acceptable in local varieties of English, such as Malaysian English (Hashim 2020, which again may facilitate switching of nouns in both directions. Additional conditions favouring CLX are situations where societal norms for language behaviour are relatively loose, in that no strong stigma is attached to code-switching, and there is a balance between the two languages (no strong competition between language communities).

Finally, the typology is new in that Muysken links it to a model of bilingual optimization strategies aimed at explaining why language contact leads to such a great variety of outcomes. The optimization strategies build in part on the work of Silva-Corvalán (1994, p. 207), who also pointed out that 'in language contact situations bilinguals develop strategies aimed at lightening the cognitive load of having to remember and use two different linguistic systems'. In Muysken's approach, insertion is seen as a strategy where the speaker uses as much as possible of the L1: in INS, the speaker only switches to the L2 for some content words, while for BFL the speaker uses as much as possible of the L2, switching back only

for some discourse markers or conjunctions. ALT is a type of code-switching that relies on universal (language-independent) principles of mixing (e.g., left- or right dislocation), while CLX is seen as a strategy aimed at matching L1 and L2 patterns where possible: the speaker fills a shared grammatical structure with content and function words from both languages.

The different types of code-switching are also relevant for models of bilingual speech processing, in that they may engage cognitive control (the ability to inhibit and monitor one's languages) to different degrees. In ALT, inhibition is arguably strongest, and in CLX least strong, with INS occupying the middle ground (Treffers-Daller 2009), although producing bilingual utterances with CLX may entail reconciling the requirements of typologically different grammars, which requires substantial monitoring skills (see Hofweber et al. 2016, et seq). It can also be assumed that the language systems are engaged in a variety of ways during code-switching. While INS and CLX rely on two languages being *simultaneously* activated in a bilingual, for ALT (and possibly BFL) they are likely activated *consecutively* (Muysken 2000). In addition, in our opinion, a key difference between INS and CLX is that for INS only *one* grammar (that of the ML) is generally activated, whereas for CLX *both* grammars actively interact. Of course, the lexicons of both languages are active too for INS as well as CLX, whereas for BFL only a small subsection of the lexicon of the first language is activated (discourse markers and some co-ordinate conjunctions) in addition to the lexicon and the grammar of the second language.

While Poplack and Walker recognize the originality of Muysken's typology, because it brings together the different types of code-switching in a new unifying framework, they also contend that CLX remains 'the weakest link' in the typology 'because its nature (and even its existence) have not been subjected to the rigors of the variationist method' (Poplack and Walker 2003, p. 682). Possibly in response to Poplack and Walker's (2003) critique, Deuchar et al. (2007) took up the challenge of providing corpus linguistic evidence for the typology in general and for the existence of CLX in particular. They carried out analyses of 300 switches, sampled from three different corpora, using a detailed list of diagnostic criteria aimed at differentiating between the three types of code-switching and first formulated in Muysken (2000). In their paper, Deuchar et al. demonstrate that CLX can be quite frequent even in language pairs which are not closely related. In the current paper we follow Deuchar et al.'s method for establishing the code-switching types in our data.

2.3. Code-Switching in Malaysia

Malaysia is one of the most multilingual countries in the world (Manan et al. 2015). At least one hundred languages are spoken in the country alongside Malay and English, with the latter used widely as the medium of instruction in schools, in industry and government. In this highly multilingual country, code-switching has become the norm in conversations among Malays, Chinese, Indians and other ethnic groups (David 2003) and David et al. (2009a) note it is so entrenched in Malaysia that it appears to have become a code in its own right. Interestingly, it is not only in informal situations, such as the home domain (David et al. 2009c) or informal discussions in schools (Ariffin 2009) or on social media (Bukhari et al. 2015; Rasdi 2016) that multilinguals code-switch. It is also quite common in more formal settings, such as classrooms at different levels of education, the court room, in organizational emails (Habil and Rafik-Galea 2009) and newspapers (see David et al. 2009b).

For the purposes of the current paper we will mainly focus on code-switching between Malay and English, albeit in the awareness that this is not the only language combination in which Malaysians code-switch, and some switch between three languages (McLellan and Nojeg 2009). Because functional aspects of code-switching (particularly of classroom discourse) have been studied in detail already in the studies mentioned above, we will concentrate here on the linguistic characteristics of Malay-English code-switching. To the best of our knowledge, there is no comprehensive overview of the linguistic characteristics of Malay-English code-switching, and the variability in the patterns. We therefore begin by

bringing together the available evidence from the available literature. After describing the code-switching patterns, we use Muysken's (2000, 2013) typology to analyse the data.

When we use the term 'code-switching' in this paper, it covers switches of multiword sequences as well as single words, some of which might be established or new borrowings in either of the two languages. The earliest sources on code-switching in Malaysia we have been able to find are N. Abdullah (1975) and Ozog (1987). N. Abdullah (1975) notes that in informal situations, 'vacillation' between Malay and English takes place among Malay-English bilinguals, and that 'constant vacillation' is likely when participants know each other well and are proficient in Malay. A slightly different view emerges from Ozog (1987), who notes the 'mixed language' is only used in intra-ethnic communication in informal situations, while in interethnic communication Malaysian English or Malay is used. As this paper was written more than 35 years ago, it seems that the situation Ozog (1987) describes has changed in that, as noticed above, code-switching now also takes place between members of different ethnic groups (David et al. 2009a) and is currently frequent in more formal domains too. However, it is possible, perhaps even likely, that code-switching is not exactly the same across all the different formal and informal domains in which information is exchanged: informal contexts, such as conversations among friends or exchanges on chat forums, or other social media are more likely to allow for unrestricted mixing than more formal situations where stricter rules for language use may exist. To the extent that the current state of research makes it possible, we will try to throw some light on the issue of the variability in code-switching patterns in Malaysia.

N. Abdullah (1975) studied code-switching among 25 Malay university students in the UK (from the Malay ethnic group), while Ozog's (1987) data were collected in 'casual conversations' in staff rooms or student accommodation common rooms at schools and universities in Malaysia (21 informants). Bukhari et al. (2015) and Rasdi (2016) used Facebook data from Malay-English bilinguals, and Wong (2012) studied code-switching among two groups of female Malaysian-Chinese bloggers (eight informants in total). The younger group consisted of 20–35 year-old females, and the older one of bloggers of 51 years old and above. The languages used by the bloggers included English, Malay, Mandarin (Chinese dialect), Japanese, Spanish, Cantonese (Chinese dialect), Hokkien (Chinese dialect) and Foochow (Chinese dialect). Another important source on variability in Malay-English code-switching is McLellan (2009b), who studied messages on online discussion fora in Brunei, a country situated on the north coast of Borneo. Finally, Majid (2019) collected data from two English language lecturers in a university in Malaysia whose classes were recorded for seven weeks.

We will first describe the range of phenomena found in different data sets. As in many code-switching data there is an asymmetry between the treatment of open class and closed class items (Joshi 1982), we have divided our presentation into these two categories. After presenting switches in open and closed class items, we will interpret these in the light of Muysken's (2000, 2013) typology.

2.3.1. Switching of Open Class Items

Nouns and Nominal Groups

In this section we will first pay attention to switches of single nouns, and will then present and discuss word order in mixed nominal groups. As is common in most code-switching data, there are many switches of single nouns in Malay-English code-switching. Most often it is English nouns that occur in Malay. N. Abdullah (1975) makes a distinction between (a) the occurrence of single English nouns in Malay utterances, for which there is no equivalent in Malay (e.g., *heater*, *central heating*, *estate*, *theory* and *practical*); (b) words which do exist in Malay but for which the Malay equivalent is less common (e.g., *summer*, *winter*, *shopping* and *machine*), and (c) words which are typical for Western cultures or which belong to the domain of technology and education (*assess*, *economics*, *law*, *psychology*, etc.). She notes there is considerable overlap between the vocabularies of English and Malay in a range of domains.

Ozog (1987) does not discuss the semantic fields to which English nouns belong but notes that English nouns are not accompanied by articles, as in (10) and (11), where the definitive article is omitted.

(10) *As you* ambil *pattern*
As you bring pattern
'As you bring the pattern.' (Ozog 1987, p. 74)

(11) I beli kat *airport*
I bought at airport
'I bought it at the airport.' (Ozog 1987, p. 74)

The omission of articles is very common in code-switching data from a wide range of language pairs, particularly when the ML of the utterance does not have articles (Myers-Scotton 2002; Owens 2005). Across the different Malay-English data sets that have been described in the literature, switching of so-called 'bare nouns' is one of the most frequent types of switches, which is likely to be related to the fact that there are no articles in Malay, as well as to the omission of articles in Malaysian English (Wong 1981).

An alternative to the use of bare nouns is the application of reduplication to English nouns in Malay utterances, as in (12), where the '2' indicates reduplication.

(12) . . . seperti *manuscript*2 atu.
like manuscript2 DEM
' . . . like those manuscripts.' (McLellan 2009b, p. 11)

English nouns that are embedded into Malay are generally not marked for number, probably because in Malay number is not marked on nouns but indicated through reduplication or a quantifier, such as *semua* 'all', as in (13a–b). McLellan (2009b), however, found fourteen examples of switches of nouns for which the English plural was retained, as in (14).

(13a) murid 'pupil'
(13b) murid-murid 'pupils' (Nadarajan 2006, p. 42)

(14) Jangan tah sabut *benefits* keraja'an Brunei
NEG-IMP DM mention benefits government Brunei
'Don't mention the benefits to the Brunei government.' (McLellan 2009b, p. 10)

Attachments of English plural morphology to a Malay noun are extremely rare in datafrom face-to-face communications: we have found only such one example, namely (15), where a plural -s is attached to *cawangan* 'branch'. According to the authors, Ariffin and Rafik-Galea (2009), this could be a form of language play, employed to enliven the conversation, because the switch was considered to be very funny by the audience, which underlines the exceptional status of this example.

(15) *There are five* cawangan*s* *here*, cawangan*s*, *ya*.
There are five branches here, branches, yeah.
'There are five branches here, branches, yeah.' (Ariffin and Rafik-Galea 2009, p. 12)

McLellan (2009b) confirms that there are no cases of English inflection (plural) attached to Malay nouns.

However, it seems that this restriction on the use of English plural does not apply to code-switching in social media. Bukhari et al. (2015) found 16 examples of English plurals attached to Malay nouns in Malay contexts, as in (16) and 24 examples of English plurals attached to Malay nouns in English contexts as in (17).

(16) Seri pengantin*s*! Nampak?
Gorgeous brides! See?
'Gorgeous brides! Do you see (them)?' (Bukhari et al. 2015, p. 7)

(17) *To all my* sayangs, *congratulations on ur C-Day(u know who u r).*
To all my dearest, congratulations on your convocation day (you know who you are)
'To all my dearest, congratulations on your convocation day (you know who you are).'
(Bukhari et al. 2015, p. 7)

It is also interesting to note that articles appear variably with inserted nouns, often depending on the ML of the clause. In examples such as (10) and (11), there are no articles accompanying the English nouns. In these utterances, Malay can be considered as the ML, if the root of the main verb is taken to be the criterion for determining the ML. By contrast, when the ML is English, articles can accompany switched nouns, as in (18), where Cantonese *kelefeh* 'an extra, an unimportant person' is preceded by an article. Article use appears to be variable, as can be seen in (19) where the main verb is English but there is no article in front of the noun.

(18) I was merely a *kelefeh* whom he thought he can take his anger on me
I was merely an unimportant person whom he thought he can take his anger on me
'I was merely an unimportant person whom he thought he could direct his anger to.'
(Wong 2012, p. 81)

According to McLellan (2009b) examples such as (19), where a Malay NP is inserted into an English clause, are much less common than the reverse. The ones that are found represent cultural items that are difficult to express with English translation equivalents.

(19) BAN *pasar malam*
Ban market night
'Ban the night market.' (McLellan 2009b, p. 431)

We will now turn to word order in mixed nominal groups. Rasdi (2016) offers different examples of compounds in which the English non-head appears after the head. In Standard English *chocolate* would appear before the head *bouquet* in (20) and *psycho* before the head *roommate* in (21).

(20) Apa salah kalau senior nak datang bertandang, dengan satu *bouquet*
What wrong if senior want come AV-visit with one bouquet
chocolate.
chocolate
'There is nothing wrong if the senior student wants to come with a chocolate bouquet.'(Rasdi 2016, p. 38)

(21) Sis nok kelik buat kek lapis doh-ni. bose ado *roommate* *psycho*
sis want return make cake layer already DEM bored have roommate psycho
'I want to go back home and make Kek Lapis (Layered Cake). I'm bored of having a psycho roommate.' (Rasdi 2016, p. 38)

The order appears to depend on the individual compound, because in other utterances, English compounds which follow English word order are found. It is possibly the ones that are relatively fixed, such as *open order*, *liquid lipstick* and *honey bee* (Rasdi 2016, p. 92), which appear in the standard British order. *Chocolate bouquet* in (20) and *psycho roommate* in (21) are not widely used compounds in English, and the internal structure of these compounds may therefore be more malleable.

More complex English modifiers can also appear after a Malay head noun. In (22) we find an English modifier consisting of an adjective and a noun, *cotton candy*, after the Malay head noun *hati* 'heart'.

(22) Ku x-sampai hati, hati *cotton candy* cepat kesian
1s NEG-reach heart heart cotton candy fast pity
'I don't have the heart to do it, my cotton candy heart pities people too easily.'
(Rasdi 2016, p. 52)

In fact, even when both the head noun and the adjective are English, word order can be Malay, as in (23):

(23) *I* ada *neighbour Indian*
I have neighbour Indian
'I have an Indian neighbour.' (Ozog 1987, p. 74)

But word order in a mixed NP does not seem to always follow Malay rules, as according to Ozog (1987), we can find an English NP functioning as a modifier in another NP (with a Malay head), as in (24), where *form three* appears between the Malay determiner *tu* 'this' and the Malay head noun *seorang* 'person'. The word order of this phrase is remarkable, because the determiner *tu* generally appears after nouns in Malay, and not at the start of the NP (Ozog 1987). It therefore seems that the word order of this NP is partly English despite the fact that the head noun and the determiner are Malay.

(24) Tu *form three* seorang
DET form three boy[3]
'A form three boy.' (Ozog 1987, p. 79)

Further evidence that word order in a mixed NP can be highly variable and sometimes follow English and sometimes Malay word order can also be obtained from McLellan (2009b).

Verbs

Apart from switches of nouns, there are many examples of switches of English verbs in Malay utterances and vice versa. In (25) the English verb *tag* appears in a Malay utterance, but interestingly, *tag* is not preceded by infinitival *to*, which would be required if this was in standard English. This grammatical morpheme is omitted in many similar constructions, as for example in (26), where there is no infinitival *to* before *compare*), just like articles are omitted in front of English nouns in code-switched utterances where Malay is the ML.

(25) Jangan lupa *tag* rakan rakan *blogger* anda
NEG-IMP forget tag friend REDP blogger 2sPOSS
'Don't forget to tag your blogger friends.' (Rasdi 2016, p. 12)

(26) Tak perlu la nak *compare* laki kau dengan. orang
NEG need DM want compare man 2s with people
'No need to compare your boyfriend with other's boyfriends.' (Rasdi 2016, p. 29)

Again, the omission is related to the fact that Malay is the ML in the clause, because in (27), where English is the ML, infinitival *to* does appear before the English verb *throw* as well as before a Malay verb *canai* 'to knead'.

(27) I found out that it's easier to throw *keliling kepala* than to *canai*
I found out that it's easier to throw around head than to knead
'I found out it's easier to throw (the dough) around the head than to knead.' (N. Abdullah 1975, p. 32)

More evidence that infinitival *to* can appear directly before a Malay verb can be found in (28), where we find that *layan* 'to entertain' is preceded by infinitival *to*.

(28) *Wasn't in a mood* *to* layan *any frickin'* *promoters but I* *took* *the handout and*
Wasn't in a mood to entertain any frickin' promoters but I took the handout and
smiled to her
smiled to her
'I wasn't in a mood to entertain any frickin' promoters but I took the handout and smiled to her.' (Wong 2012, p. 69)

Switches of verbs may also consist of verb-particle combinations, as can be seen in *give up* in (29).

(29) *I*, lepas dua tiga biji *give up* la!
I after two three seeds give up DM
'After two or three seeds, I give up!' (N. Abdullah 1975, appendix)

Switches of verbs are particularly interesting because verbs impose clear selection restrictions on their environment and can carry inflections which are used to establish relationships between constituents in a sentence. Because of the complex syntactic relationships between verbs and other constituents in a sentence, in many language contact situations, alien verbs receive special treatment before they can be inserted into a language. However, Muysken (2000, p. 185) notes that languages which lack verbal inflection can incorporate verbs from another language without further adaptation. As Malay verbs are not inflected for tense (McLellan 2009b), Malay probably belongs in the group of languages which can easily incorporate English verbs. While English verbs can carry inflections, the English verbs which appear in Malay utterances are all non-finite, which may have facilitated the switch. Most verbs are in the infinitive form, but some are in the *-ing* form, as in (30), where *posing* appears after the Malay modal verb *kena* 'must'. In standard English an infinitive form would have been expected after a modal verb.

(30) Sbb[4] tu kena *posing* silang kaki. Haha!
Because DEM must posing cross leg. Haha!
'That's why I have to pose with my legs crossed. Haha!' (Rasdi 2016, p. 41)

McLellan (2009b) also notes that switches of Malay verbs in English utterances are less common than switches of English verbs in Malay. An example is given in (31). Importantly, McLellan points out that there are no cases of Malay verbs with English bound morphemes.

(31) *Then at the end of time our population* jadi *0*
Then at the end of time our population become zero
'Then at the end of time our population will become zero.' (McLellan 2009b, p. 13)

As Ozog (1987) notes that English 'verbal groups' are almost always part of a wholly English clause, and examples such as (30)–(31) are not present in his overview, it is possible that switches of lone English non-finite verbs in Malay utterances constitute innovations by comparison with the data from the 20th century.

Adjectives

Switches of adjectives are fairly rare but an example can be found in (32), where the Malay adjective *comot* 'stained' appears before the noun *track pants*.

(32) *The other is me, dressed in my tattered t-shirt and* comot *track pants, sans*
The other is me, dressed in my tattered t-shirt and stained track pants, without
make-up, uncombed hair . . .
make-up, uncombed hair . . .
'The other is me, dressed in my tattered t-shirt and stained track pants, without
make-up, uncombed hair and stained track pants . . . ' (Wong 2012, p. 55)

The fact that the Malay adjective appears *before* the noun shows clearly that English grammar rules apply in this noun phrase, because adjectives normally appear after the noun in Malay (McLellan 2009b). The opposite can be seen in sentences where Malay is the ML. In (33), for example, an English adjective appears in a Malay NP, after the noun. Thus, *best*[5] follows *adengan* 'scene'.

(33) . . . .sebab aku tau mesti ada adengan *best* untuk aku tengok.
. . . because I know must have scene best for 1s watch
' . . . because I know there must be some exciting scenes for me to watch.' (Rasdi 2016, p. 43)

English adjectives can also be used predicatively in Malay, as in (34), where *boring* is the predicate in an utterance where Malay is clearly the ML. According to Ariffin and Rafik-Galea (2009), *boring* can express boredom or dislike, and the latter is the intended meaning in this example.

(34) Saya *boring* betul kalau benda-benda jadi macam ni
I boring Very if things happen like this
'I really don't like it when these things happened.' (Ariffin and Rafik-Galea 2009, p. 11)

When used predicatively, a degree adverb may appear after the English adjective, as in (35). The word order in the AP is clearly Malay here.

(35) . . . so takde la nampak *plain* sangat.
. . . so NEG-have DM look plain very
' . . . so that it won't look very plain.' (Rasdi 2016, p. 44)

Adverbs and Discourse Markers

Single Malay or English adverbs (temporal, manner and locative) or adverbial groups are frequently switched (Ozog 1987), because these are adjuncts and there are few if any restrictions on switching this type of expression (see also Treffers-Daller 1994). An example is given in (36) where *from now on* is used at the start of the Malay utterance. Providing a complete list of all the different adverbial expressions that occur in Malay-English code-switching is beyond the scope of the current project.

(36) *From now on*, sila *whatsapp number* baru mek ye untuk urusan sebarang
From now on please whatsapp number new 1s DM for any business
'From now on, please Whatsapp me on my new number for any business'
(Rasdi 2016, p. 13)

Switches of a degree adverb, such as *lagi* 'repeat, much, more' in (37), are much less common in other language contact situations.

(37) *This is grandma* *she then* *helped me* *wear this* lagi *elaborated* *kimono*
This is grandma she then helped me wear this very more elaborated kimono
'This is grandma, she then helped me wear this much more elaborated kimono.'
(Wong 2012, p. 89)

At the end of a clause, one often finds Malay discourse markers, the most frequent of which is *-lah*. According to Ozog (1987) it can occur in clauses that are entirely in English, entirely in Malay or mixed (see (38) and (39)).

(38) *staff room* ni bising lah
Staff room this noisy DM
'The staff room is noisy.' (Ozog 1987, p. 87)

(39) *you understand* lah kan
You understand DM DM
'You understand, don't you?' (Ozog 1987, p. 87)

While the specific meanings of discourse markers are often difficult to capture, it is possible to evaluate their pragmatic functions. Tay et al. (2016) suggest *lah* modifies the utterance from one that has a highly assertive tone to a polite request or a friendly encouragement.

Idioms and Fixed Expressions

N. Abdullah (1975) notes the use of English idioms and fixed expressions such as *by the way* in (40) that occur in Malay utterances.

(40) *By the way*, bila Agung nak bagi *you* Tun?
By the way, when Agung want give you Tun?
'By the way, when is the Agung (King of Malaysia) going to confer on you the title of Tun?' (N. Abdullah 1975, p. 31)

Similar examples of switches of English fixed expressions can be found in (41), where *stay up* is inserted into a Malay clause.

(41) Alhamdulillah rezeki.. mlm[6] ni *stay up* lg[7] *baking* sampai subuh
God thank luck . . . tonight this stay up very baking until dawn
'Thank God, luck . . . tonight I'm going to stay up again to bake until dawn.' (Rasdi 2016)

2.3.2. Closed Class Items

Pronouns

In most data sets examples are found of the usage of the English pronouns *I* and *you* in Malay utterances, which N. Abdullah (1975) interprets as a strategy of neutrality, because of the complexities of the pronominal system in Malay with six different levels (see also Othman 2006). According to N. Abdullah (1975), who notes that English pronouns are very frequent in Malay, the choice of English pronouns makes it possible to avoid any reference to respect, seniority or power, and thus the speaker can avoid making any Malay pronoun choices that might offend the hearer (42)–(43).

(42) *I* teringat dulu *in my student days*
I remember before in my student days
'I remembered in my students days' (N. Abdullah 1975, p. 21)

(43) *You* bubuh satu *heater* dibawah
You put one heater below/down
'You put one heater down.' (N. Abdullah 1975, p. 21)

In some cases English pronouns are accompanied by *punya* 'own', as in the conversation between A and B in (44), but this usage is not discussed in N. Abdullah's work.

(44) A: Inikah *I* punya?
B: *you* punya, *you* punya
A: This+DM *I* own?
B: you own, you own
A: 'Is this mine?'
B: 'Yours, yours' (N. Abdullah 1975, p. 21)

Ozog (1987) provides several examples where *punya* appears between the modifier and the head noun, and the determiner *itu* appears in final position (45).

(45) *Baby* punya *nest* itu
Baby own nest that
'The baby's nest' (Ozog 1987, p. 79)

McLellan (2009b) and Rasdi (2016) note that there are many switches of English pronouns in their data. *You* is abbreviated to *u*, and sometimes *I* and *U* are written in lower case, as in (46). According to Rasdi, the popularity of the English pronouns can be ascribed to the efficiency of typing a single letter than to write the corresponding Malay pronouns *saya 'I'* and *awak 'you'*. Also to be noticed is the absence of the preposition *of* before *u*. This means the subcategorization frame for *mimpi* 'dream' is Malay. Assuming *mimpi* sets the grammatical frame, English grammatical morphemes are not expected in this clause.

(46) Mlm td ***i***[8] mimpi ***u*** t au.
Night recently dream you DM
'Last night I dreamt of you, you know' (Rasdi 2016, p. 47)

For the third person singular, no English pronouns are used. It is possible that English pronouns of the third person are less popular because they distinguish between males and females, which is not the case for the Malay pronouns of the first and second person. In addition, there are fewer levels to choose from for pronouns of the third person: *dia* 's/he' is the low variant and *beliau* 's/he' high variant, which means the choice is less complex than for the first and second persons. In addition, the most direct threat to a person's face is likely to come from making a mistake in the choice of forms of address in the presence of the other interlocutor(s), which need to be addressed with pronouns of the second person (not the third). Those we are talking about are not likely to be present in the conversation, which diminishes the risk of embarrassment.

The fact that pronouns are switched so frequently is remarkable because of the restrictions on switching of function words, which was already noticed by Lehtinen (1966, p. 177), who also notes that there could be exceptions to this constraint 'in cases where such a switch is forced by structural considerations.' It seems that the pragmatic reasons mentioned above constitute relevant structural considerations which allow the speaker to overrule the constraint against switching of function words. It may also be of interest to note that the use of English pronouns in Malay is mimicked in movies, as shown in Nil and Paramasivam (2012), who analysed conversations in *Gol dan Gincu* '*Goal and Lipstick*', and portrays the lifestyle of youths living in a college in Kuala Lumpur.

Ozog (1987) notes that subject forms of the English pronouns are often used as direct objects as in (47), indirect objects as in (48) or as possessives in (49). According to the author this was at the time of writing much more common than the use of *me* as an (in)direct object form.

(47) dia *stop I*
She stop me
'She stopped me.' (Ozog 1987, p. 75)

(48) *You* tak tulis bagi *I* resit ke
You not write give me receipt DM
'You didn't write and give me a receipt, did you?' (Ozog 1987, p. 75)

(49) *Ruler I*
Ruler my
'My ruler.' (Ozog 1987, p. 76)

It needs to be clarified whether the usage of subject forms of the pronouns *I* and *you* for a variety of functions as illustrated in (47) to (49) is still common in Malaysian-English code-switching, as further examples from more recent data sets do not contain this pattern. In the more recent data we have access to, *I* is used only as a subject, while for *you* the subject and (in)direct object forms are the same.

Demonstratives

Apart from the personal pronouns, there are also Malay demonstratives in English utterances, namely *(i)tu* 'that', an *(i)ni* 'this', as in (50) and (51).

(50) kunci saya pada *cupboard* tu
Keys I on cupboard that
'My keys are on the cupboard.' (Ozog 1987, p. 79)

(51) *pattern* ni
pattern this
'This pattern' (Ozog 1987, p. 80)

Ozog notes these demonstratives sometimes function as a determiner and sometimes as a discourse marker, and suggests that these may fulfil the same role in the phrase as *lah* at the clause level, and functions as a marker of rapport, solidarity, informality, etc., as in

(3) given in Section 2.1. The order in which they appear is remarkable because they appear after the noun, which means the word order is clearly Malay in these NPs.

The Aspectual Marker Dah (Sudah) 'Already'

The Malay aspectual marker *dah (sudah)* 'already', which indicates that an action has already been completed (Sneddon 2007) is frequently found before switches of English verbs, as in (52). This is common in all data sets.

(52) dah *confirm*
already confirmed
'It has been confirmed.'(Ozog 1987, p. 81)

In some cases, *dah* appears after the verb, as in (53), but the preverbal position is more frequent.

(53) Masak ape lagi kite lemang semua *settle* dah
Cook else again see lemang all settle already
'What else is there to cook? We've cooked lemang (a rice and coconut dish) and everything else is already settled.' (Rasdi 2016, p. 77)

Modal Verbs

All data sets contain examples of switches of modal verbs, such as *boleh* 'can', indicating ability or permission (I.H. Abdullah 1993) or *mesti* 'must'. They often occur in utterances that consist of English words only, except for this modal verb, as in (54), or in Malay clauses where the modal is found just before an English verb, as in (55).

(54) *You* boleh *attracted to it*
You can attracted to it
'You can be attracted to it.' (Ozog 1987, p. 82)

(55) Borang tu mesti *sign*
form this must sign
'You must sign the form.' (Ozog 1987, p. 82)

Percillier (2016) shows that *kena* can also be used in sentences with code-switching. While *kena* can function as a modal verb, as in (30), where it means 'must', in (56) and (57) it fulfills the role of a passive marker (see also Karim et al. 2008).

(56) *He* kena *sabotage*
He was sabotage
'He was sabotaged.' (Percillier 2016, p. 20)

(57) *Confirm* kena *crush* dengan *other people* . . .
confirm get crush by other people . . .
'Confirmed (we will) get crushed by other people . . . '
(https://www.english-corpora.org/glowbe/) [accessed on 20 October 2021]

Interestingly, in (58) and (59), which were retrieved in an online corpus of Malaysian English, it is used in combination with nouns (*punishment* and *jailbreak*) rather than verbs. There is therefore a variety of structures in which it can be used, which makes this is very versatile tool for expressing grammatical relations.

(58) . . . *even those above 7–8 years old* kena *punishment* . . .
. . . even those above 7–8 years old get punishment . . .
' . . . even those above 7–8 years old got punished . . . '
(https://www.english-corpora.org/glowbe/) [accessed on 20 October 2021]

(59) . . . *the main point is*, kena *jailbreak* dulu!
. . . the main point is, have to jailbreak first!
'The main point is, we have to break out of jail first!'
(https://www.english-corpora.org/glowbe/) [accessed on 20 October 2021]

Furthermore, *hendak* 'want', *nak* 'want' and *perlu* 'need' can be used in combination with English verbs in the infinitive form, as in (60).

(60) Kita perlu *collect*
we need collect
'We need to collect' (Ozog 1987, p. 82)

Negation

The regular way to negate verbs or adjectives is by putting *tak (tidak)* 'no(t)' before the verb or the adjective that is negated (Kroeger 2014). This is also the case when the verb or adjective are in English, as in (61) or (62).

(61) *I* tak *order* (verb)
I not order
'I didn't order.' (Ozog 1987, p. 82)

(62) tak *equip* (adjective)
not equipped
'not equipped.' (Ozog 1987, p. 82)

Conjunctions

Co-ordinate conjunctions appear to be switched frequently in both directions, such as Malay *tapi* 'but' in (63), and conversely, *but* and *and* in Malay utterances, as in (64).

(63) tapi *I can understand the theory*
but I can understand the theory
'But I can understand the theory.' (N. Abdullah 1975, appendix).

(64) dah lengkap kasut, baju *and* tudung . . .
already complete shoes dress and shawl . . .
'Now that the shoes, dress and shawl are complete . . . ' (Rasdi 2016, p. 44).

Ozog (1987) notes that conditional clauses in mixed utterances can begin with the conjunction *kalau* 'if', and suggests this is much more common than starting a conditional clause with English *if*. An example of *kalau* in a Malay utterance can be found in (20) in Section 2.3.1 and in a mixed utterance in (76) in Section 3.2.2.

Prepositions

Switches of single prepositions are very rare. The only example we have found is (65), where *kat* (dekat) 'near' is used before an English placename. Because of the lack of congruence between the ways in which location and movement through space is expressed in different languages (Talmy 2000), it can be very difficult to switch for a single locative preposition. As there is no Malay equivalent for the placename Chalburn, one could of course argue that *Chalburn* is the Malay word for this place and the switch in (65) is in fact a switch of a full PP.

(65) *We are going to stay with an Englishman who owns a mansion* kat *Chalburn*
We are going to stay with an Englishman who owns a mansion near Chalburn
'We are going to stay with an Englishman who owns a mansion near Chalburn (N. Abdullah 1975, p. line 41 in appendix)

However, Sebba (1998) offers one example of a Malay preposition switched on its own in an English utterance (see 66).

(66) *I drove* sampai *the day before*
I drove until the day before
'I drove until the day before.' (Sebba 1998, p. 15)

Word-Internal Switches

The only examples we have been able to find are from two studies of code-switching on Facebook. Bukhari et al. (2015) report cases of the attachment of English gerundive suffix *-ing* to Malay verb roots, as in (67), where it is attached to *merindu* 'miss' and (68) where it is combined with the Malay verb *tiru* 'copy'. Bukhari et al. note that this happens in sentences that mainly consist of Malay words, as in (67) as well as in sentences that consist entirely of English words, as in (68).

(67) Merindu-*ing* *him* inimlm[9]!
missing him tonight!
'I am missing him tonight!' (Bukhari et al. 2015, p. 7)

(68) *I am not* tiru-*ing you*!
I am not copying you!
'I am not copying you!' (Bukhari et al. 2015, p. 7)

Rasdi (2016) contains a number of exceptional word-internal switches which have not been reported in other sources. The first of these is seen in (69), where the English derivational suffix *-ness* is attached to a Malay adjective, and in (70), where the Malay third person possessive affix *-nya* is attached to an English adjective *rare*, while in (71) it is affixed to a noun.

(69) *Haha! And you can't brain* *the* sedap*ness too*.
Haha! And you can't handle the deliciousness too.
'Haha! And you can't handle the deliciousness either.' (Rasdi 2016, p. 32)

(70) *Wowwww* *rare*-nya *wish* ni!
Wow rare-POSS wish this
'Wow, this wish is so rare!' (Rasdi 2016, p. 32)

(71) Bahawa kreativiti yang menjadi kekal *feature*-nya
That creativity that become established feature-3POSS
'That creativity that has become established as a feature.' (McLellan 2009a, p. 270).

2.4. Analyses Using Muysken's (2000) Typology

Wong (2012) and Rasdi (2016) are among the very few who used Muysken's (2000) typology in their analyses of Malay-English code-switching. Wong (2012) found that CLX was the most frequent pattern in the younger group of female bloggers in her study, while the older group engaged more in ALT and INS. The data are very interesting and provide clear evidence of the creativity of the bloggers, although it seems that in this paper the different types of code-switching distinguished by Muysken (2000) have been interpreted in a way that is rather different from the original, which makes it difficult to compare the quantitative results with those in other studies. An example of ALT from this data set can be found in (72). The switch takes place at a major clause boundary[10] between the main clause, which is in English and the subordinate clause, which is in Malay.

(72) *My husband helps me to* pergi pasar beli sayur, daging dan ikan.
My husband helps me to go market buy vegetables, meat and fish.
'My husband helps me to go to the market to buy vegetables, meat and fish.'
(Wong 2012, p. 65, from a 52 year-old blogger)

However, some of the switches which have been classified as INS seem to satisfy the criteria for BFL more than those for INS, as the switches consist of a single pragmatic particle from Cantonese (e.g., *wei*, in (73), which is attached to the periphery of an English sentence. The particle is difficult to translate but according to the author it denotes sarcasm.

(73) *While this Airasia is insane. Always give people heart attack* wei!
While this Airasia is insane. Always give people heart attack DM!
'Meanwhile Airasia is insane. They always give people heart attack!'
(Wong 2012, p. 66, from a 27-year-old female blogger)

Similar English utterances, but this time with the Cantonese particle *hor*, which signals attention, support or agreement, as in (74), can be found in these bloggers' data set (see also Tay et al. 2016 for details of pragmatic particles in Malaysian English). In addition, there are examples in Wong (2012) of other Cantonese particles (*lor, nah, aiyo*), denoting a variety of pragmatic meanings that are attached to the periphery of English utterances, all of which are probably to be analysed as BFL.

(74) *Plus* hor, *I tell you a secret, I actually made up the stories*
Plus DM, I tell you a secret, I actually made up the stories
'Plus, I tell you a secret, I actually made up the stories.' (Wong 2012, p. 67, from Lilian, a 50-year-old female blogger)

Additionally, several examples of CLX in the data could be seen as INS. In (28), for example, from a 27-year-old female blogger, the switch to Malay for *layan* 'entertain' is likely to be INS as it is a switch of a single content word, and the words before and after the switch are grammatically related.

While Rasdi (2016) only analysed ALT and INS in her data, several of her Facebook examples seem to have characteristics of CLX because a shared grammatical frame appears to have been created, which facilitates code-switching at points where no 'objective' congruence exists between both languages. In (75), for example, there is a switch to English between the Malay adjective *pandai* 'smart' and its English complement *create infographic*. In English, the complement of the adjective *good* would have to be *at creating infographic*. It seems the English part of the sentence has been adjusted to fit Malay grammar, or that the English grammar has been 'suspended'.

(75) Aa sesiapa dekat sini yang pandai *create infographic*?
Ah anybody near here REL smart create infographic
'Ah, Is there anybody here who is good at creating infographic?' (Rasdi 2016, p. 12)

The adaptation of (parts of) English sentences to Malay grammar was already noted by Ozog (1987, p. 73), who suggests that Malay is the 'dominant' language in that 'Malay syntactic patterns are often carried over into English'. However, it is also possible to go one step further and to see examples such as (75) as CLX, in that the speaker creates congruence between the grammars by drawing on the grammar rules as well as the lexicons from both languages. In doing this, the speaker facilitates switching at switch points where there is no 'objective' congruence between the languages, as argued by Sebba (1998).

The data we have seen reported in the literature appear to show that ALT is the least common type of code-switching in Malaysia. Of course participants can switch *between* utterances, which is akin to ALT, but switches *within* utterances are more likely to be INS, CLX or BFL, because switches are generally for just one word, including compounds. Only adverbial phrases or discourse markers, which are often found at the periphery of the utterance would qualify as ALT. While one utterance can contain multiple switches to Malay (or to English), these switches generally consist of relatively short expressions. Switches for longer stretches of speech, at major constituent boundaries, such as the switch in in (72) between main and subordinate clause, are rare.

It is of course possible that there is a great deal of interindividual variation that cannot be clearly captured on the basis of the evidence reviewed here. Furthermore, different code-switching conventions might apply to social media and to face-to-face conversations. It is difficult to say much about these conventions at the moment, although it appears that switching within words is only found in social media. In fact, so far only one word-internal switch (15) of an English plural attached to a Malay noun has been found in a data set from face-to-face conversations. It is possible that the informality of much of the interaction on social media, such as Facebook or blogging sites, is most conducive to creative forms of code-switching. It could also be that societal norms for 'correct language use' are somewhat relaxed on social media, and that the remoteness of such exchanges reduces the chances of embarrassing oneself and losing face. All this might lead to creative forms on social media that are not found elsewhere. To what extent these assumptions are correct will need to be studied in future research.

3. The Current Study

3.1. Aims and Research Questions

As explained in Section 2, there are very few studies in which Muysken's (2000) typology has been used to analyse code-switching in Malaysia, and to the best of our knowledge, none which have used his (2013) model. We aim to contribute to testing the model and developing it further by focusing in particular on one of the most controversial types of code-switching in the model, namely CLX. Because switches of function words are among the distinctive features for CLX, and there are many switches of Malay function words in English utterances in the literature, data from Malay-English code-switching are particularly interesting for testing the model. As explained in the introduction, CLX may be expected in this language pair because there is a long tradition of language contact, with evidence for convergence between all different languages spoken in Malaysia (Vollmann and Soon 2020), and there appears to be less stigma attached to code-switching than in other contact situations (David et al. 2009a).

Our study builds on the work of Deuchar et al. (2007), who demonstrated how the distinctive features for INS, ALT and CLX can be used to determine which type of code-switching a particular instance of code-switching is likely to represent, and which type of code-switching is the dominant one in a corpus. This example has so far only been followed by Hofweber et al. (2016) for German-English code-switching. After providing an overview of the data from our corpus, we present the Malay function words which are switched in our data, and discuss whether they are indeed examples of CLX. Subsequently, we discuss the similarities and differences between the use of these function words in our data and those found in the literature. We hope these analyses will shed some light on why function words are such a specific target in our paper.

Our research questions are as follows:

(1) Which types of code-switching are found in Majid's (2019) corpus of teacher language, and which of the four types are the dominant ones in the corpus?
(2) How diverse is code-switching in the data set? To which syntactic categories or word groups do the switches belong, and how frequent are these?
(3) What are the similarities and differences between the types of code-switching found in Majid's (2019) corpus and the corpora discussed in the literature review? In which directions does switching take place most frequently?
(4) To what extent do Malay function words in English satisfy the criteria for CLX?
(5) To what extent can bilingual optimization strategies explain the popularity of switches of function words in Malay-English code-switching?

3.2. Methods

3.2.1. Participants

Participants were a male (Ali) and a female (Azma) English language lecturer from a university in Malaysia. Both were 31 years-old at the time of recording. They were

bilingual in Malay and English, with Malay as their first language and English as their second language. Both had studied for university degrees in Malaysia and obtained a Masters in TESOL. They were both fluent in English, although their English language levels were not measured. Azma taught reading while Ali taught an academic writing class. The students who were enrolled at the university where the data collection took place were also bilingual speakers of Malay and English. They were 19–20 years old at the time the data collection was carried out. Their language was not analysed for the purposes of this project, as ethical approval for using their data could not be obtained. It would have been preferable to obtain data from more teachers, but this was not possible in the scope of the study. While we cannot claim that the speech samples we obtained are representative for classrooms across similar contexts in Malaysia, it is likely they provide a good impression of the kinds of code-switching used by Malay-English bilingual teachers.

3.2.2. Instruments and Data Analysis

A total of seven classroom sessions (15 h) were video recorded for each participant at a university in Negeri Sembilan, a state 70 km south of Kuala Lumpur. The recordings were then transcribed in CHAT (Codes for the Human Analysis of Transcripts) and analysed with CLAN (Computerised Language Analysis (CLAN), developed by MacWhinney (2000). The characteristics of each utterance which contained code-switching were analysed following Deuchar et al.'s (2007) example, albeit with the addition of BFL which was not part of Muysken's typology at the time. Subsequently, the second author coded all the data, and compared her results with those of another researcher, who was familiar with the typology and coded sixteen utterances with code-switching too. Any discrepancies were resolved through discussion among the research team. After identification of the type of code switching each switch was tagged as either [@ins] for insertion, [@alt] for alternation, [@clx] for Congruent Lexicalization and [@bfl] for back flagging. After completion of the coding, the frequency of each type of intra-sentential code switching was calculated using CLAN.

As Deuchar et al. (2007), point out, it is not easy to determine which elements are switched in a sentence with code mixing, particularly if speakers switch back and forth many times within an utterance as in (76), in which Azma starts in Malay, then switches multiple times to English. It then becomes very difficult to say whether the speaker uses English words in Malay or vice versa. Determining the ML is equally difficult.

(76) Macam mana *moderator* nak *re-cap* kalau *moderator* tak faham
How moderator want re-cap if moderator not understand
what the article is about?
what the article is about
'How does the moderator want to recap if the moderator does not understand what the article is about?'

Muysken (2000) and Deuchar et al. (2007) consider different criteria that can be used to determine the ML. While an attractive structural option is to consider the highest element in the syntactic tree (i.e., the inflection on the finite verb) as indicating the ML (Klavans 1985; Treffers-Daller 1994), this is difficult in Malay because of the absence of overt finiteness markers on verbs. In the absence of overt inflection, an alternative could be to consider the language of root of this finite verb as the ML. This would mean that Malay is the ML in (76). Following Myers Scotton's logic, this would then explain why in (76) all function words in the main clause (*macam mana* 'how', *kalau* 'if' and *tak* 'not') all come from Malay. The remaining words in this clause are English content words, which could be seen as insertions into this frame. Only for the last part of the utterance, is there a complete switch to English, with the embedded clause *What the article is about*, where the inflected verb *is* sets the grammatical frame. The telegraphic style of the main clause, without any English function words, also makes it unlikely that English is the ML.

However, this does not work for all utterances. In (77), where *tahu* 'know' is the main verb, Malay could be seen to set the grammatical frame, but the presence of English function words (the pronoun *you* and the article *the*) is then unexpected. For the relative clause, *bagi* 'give' is the main verb. If this is taken to set the grammatical frame, the presence of the Malay relative pronoun *yang* 'that' is expected, but the presence of an English article before *writer* or justification is not. Given the fact that open class as well as function words are drawn from both languages in both parts of the utterance, a more likely analysis is to say that (77) is an example of CLX, also because the grammatical structure is shared between both languages.

(77)

and relevance	*here*	*you*	sendiri	pun tahu	*relevant*	ke tak	*the justification*
and relevance	here	you	alone	also know	relevant	or not	the justification
yang *the writer*	bagi						
that the writer	give.						

'And relevance here you yourself know whether the justification that the writer give is relevant or not.'

Similarly in (78), the Malay verb *tanya* 'ask' is the main verb of the clause, but all function words come from English. Alternative options are to count the number of morphemes in each language. Then, English would be the ML, and *tanya* would be inserted into this matrix. However, if many different criteria are being used to identify a ML, the concept of the ML loses its attractiveness. A more likely option in cases such as (78) is to consider these as CLX, for which the grammatical frame is shared, and there is no ML.

(78)

At least you tanya
At least you ask

'At least you ask.'

Table 2 shows how we have used Muysken's criteria for the analysis of Malay-English code-switching. We follow the method developed by Deuchar et al. (2007) but have added the features for BFL from Muysken (2014) because in 2007 backflagging was not part of Muysken's model.

Columns 6 to 10 in Table 2 shows how the criteria are applied to an example from Majid's (2019) Malay-English code-switching corpus, where a single English noun (*justification*) is inserted into a Malay utterance. In columns 1–5 we present the distinctive features from Table 1 (in grey). Column 6 shows whether a feature applies to the switch of *justification*. Thus, for example, a+ is given for single constituent and a—for morphological integration in column 6, because *justification* is a single constituent and it is not morphologically integrated into Malay. In columns 7 to 10 we compare the scores from column 6 against the criteria for that feature as given in columns 2 to 5. As explained in Deuchar et al. (2007), a score of 1 means that the feature value is as expected for that type, while −1 indicates that it is the opposite of the expected value and a score of 0 that it is neutral with respect to that feature. Thus, there is a score of 1 for INS for the feature 'single constituent' in column 7, but a 0 for ALT, because alternation does not necessarily involve switches of single constituents. For the feature morphological integration, we have given 0 for *justification*, because there are no inflections or derivational suffixes that could be attached to *justification* to integrate it into Malay. In the last row of Table 2, the scores from columns 7–10 are added up. This row shows that the highest scores are obtained for INS because 16 features are indicative of INS, and 2 counter-indicative, which means a total score of 14 in favour of INS. The other types of code-switching obtain far lower scores. For ALT, the result is clearly counter-indicative, with a negative score of −12.

After the explanation of the ways in which we analysed the data, we now turn to the analysis of the code-switching patterns in the utterances from the two teachers.

Table 2. Muysken's (2014) features applied to Malay-English code-switching.

Features Distinguishing the Four Code-Switching Strategies (Muysken 2014)					**Ada *Justification* Tak Dalam tu?** **There Is Justification Not in That.** **'Is There No Justification in There?'**				
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Diagnostic feature	Insertion (INS)	Alternation (ALT)	Congruent Lexicalization (CLX)	Back-Flagging (BFL)	scores	INS	ALT	CLX	BFL
CONSTITUENCY									
Single constituent	+	0	0	+	+	1	0	0	1
Several constituents	−	+	0	−	−	1	−1	0	1
Non-constituent	−	−	+	−	−	1	1	−1	1
Nested aba	+	−	0	−	+	1	−1	0	−1
Non-nested aba	−	+	+	+	−	1	−1	−1	−1
ELEMENT SWITCHED									
Diverse switches	−	0	+	−	Type: noun				
Long constituent	−	+	−	−	−	1	−1	1	1
Complex constituent	−	+	−	−	−	1	−1	1	1
Content word	+	−	−	−	+	1	−1	−1	−1
Function word	−	−	+	−	−	1	1	−1	1
Adverb, conjunction	−	+	−	+	−	1	−1	1	−1
Selected element	+	−	+	−	+	1	−1	1	−1
Emblematic or tag	−	+	0	+	0	0	0	0	0
SWITCH SITE									
Major clause boundary	0	+	0	+	−	0	−1	0	−1
peripheral	0	+	0	+	−	0	−1	0	−1
Embedding in discourse	0	+	0	+	0	0	0	0	0
Flagging	−	+	−	0	−	1	−1	1	0
Dummy word insertion	+	0	−	−	−	−1	0	1	1
Bidirectional switching	−	+	+	+	Switch into English				
PROPERTIES									
Linear equivalence	0	+	+	+	−	−1	−1	−1	−1
Telegraphic mixing	+	−	−	−	+	1	−1	−1	−1
Morphological integration	+	−	+	−	−	0	0	0	0
Doubling	−	+	−	−	−	1	−1	1	1
Triggering	0	0	+	−	−	1	0	0	1
Mixed collocations	0	−	+	−	−	0	1	−1	1
Self-corrections	−	+	−	0	−	1	−1	1	0
					Total score	16 − 2 = 14	3 − 15 = −12	8 − 7 = 1	10 − 9 = 1

Features characteristic for a particular code-switching strategy are marked with a +, features that are counter-indicative for that strategy are marked with −. If a feature is neutral for that strategy, it is marked with 0.

4. Results

Our first research question asked which types of code-switching were represented in our corpus and which of these were dominant in the data. In this section we will first give an overview of the quantitative results for the entire data set. In the following sections we will look at the diversity and the directionality of the switches (research question 2), and then we will focus in detail on the similarities and differences between our data and those from other sources (research question 3). After this, we discuss whether or not there is evidence that the switches of function words constitute examples of CLX (research question

4), and we finish with an interpretation of the results in the light of Muysken's optimization strategies (research question 5).

4.1. Frequency of Different Types of Code-Switching

The results for each type of code-switching are presented in Table 3. In total 1044 utterances with intrasentential switches were recorded, of which 406 were found in Azma's and 638 switches in Ali's speech. The total number of switches is higher (1217), as many sentences contain more than one switch. Table 3 shows that there is considerable variation in the code-switching strategies used by both teachers. Azma uses CLX and ALT most frequently, while for Ali it is INS and CLX. For both teachers BFL is the strategy they favour least. The differences between the teachers seem to indicate that there is considerable variation between speakers in their use of the four different code-switching strategies. Further research on larger datasets is needed to provide further evidence about the frequency of the four strategies in different contexts.

Table 3. Frequency of the four types of code-switching in our data set.

Participants	INS		CLX		ALT		BFL		Total	
	n	%	*n*	%	*n*	%	*n*	%	*n*	%
Azma	71	17.5	144	35.5	132	32.5	59	14.5	406	100
Ali	273	42.8	176	27.6	127	19.9	62	9.7	638	100
Total	344	33	320	30.7	259	24.8	121	11.6	1044	100

4.2. Diversity of Switches

Our second research question focused on the diversity of the switches in our corpus. As can be seen in Table 4, the data set is very diverse in that many function words (in particular personal pronouns, but also modal verbs, demonstratives and relative pronouns) are switched, which calls for an explanation.

Table 4. Overview of syntactic categories of the switches.

Syntax	*f*	%
Noun	261	21.50%
NP (noun phrase)	131	10.80%
Personal Pronoun	127	10.40%
Discourse Marker	121	9.90%
Lexical Verb	100	8.20%
Conjunction	95	7.80%
Adv & AdvP	83	6.80%
VP	81	6.70%
Adj & AdjP	47	3.90%
WH-Word	41	3.40%
Prep and PP	24	2.00%
Demonstrative	21	1.70%
Modal Verb	17	1.40%
Various Constituents	14	1.20%
Negation	13	1.10%
Relative Clause	10	0.80%
Possessive Pronoun	7	0.60%
Auxiliary verb	6	0.50%
Complement Clause	5	0.40%
Interjection	4	0.30%
Conditional Clause	3	0.20%
Relative Pronoun	3	0.20%
Main Clause	3	0.20%
Total	1217	100

Note: *f* = frequency, % = percentage.

4.3. Similarities and Differences between Data Sets

Our third research question addresses the issue of the similarities and differences between the data sets differ per syntactic category. We begin the overview with the content words, after which we look at the different categories of function words in turn.

4.3.1. Open Class Items

As in many other code-switching corpora, there is a large proportion of switches of nouns or nominal groups, such as *written article analysis* in (79).

(79) So jangan bagi alasan kata tak sempat buat *written article analysis*
So don't give reason why not have time make written article analysis
sebab ada dua *submission*.
because there are two submission
'So don't give the reason why you didn't have time to make a written article analysis because there are two submissions.'

There is a clear asymmetry in the direction of switches of nouns: the large majority of these switches (290) consist of English nouns in Malay utterances, and only about 70 of these consist of Malay nouns in English utterances. This is likely to be related to the fact that the data were collected in the context of an English language classroom. Therefore many of the nouns/nominal groups are expressions related to English language and linguistics, as in (79). However, there is one Malay expression, *maksud-nya* 'meaning-POSS, which means', that occurs in mixed utterances in situations in which lecturers explain the content of a class to their students. In (80) the lecturer wanted to explain the meaning of *wasteful* to the students by relating it to the root word *waste*.

(80) Kalau dia *wasteful* maksud-nya dia tak *waste*?
If s/he wasteful mean-POSS s/he not waste?
'If he/she is wasteful does it mean that he/she does not waste?'

The expression *contoh-nya* 'for example' also occurs very frequently in mixed utterances such as (81), where its role appears to be to flag that an explanation is coming.

(81) Contohnya, *there are several types of pollution which is a,b,c.*
For example, there are several types of pollution, which is a, b, c.
'There are several types of pollution, for example a, b and c.'

These two forms of code-switching have not been found in other sources on code-switching in Malaysia, possibly because explanations are typical for classroom settings.

We have found one example where an English plural form was used in a Malay structure, see (82). Interestingly, the English word *students* was also reduplicated, which means plural was expressed twice: first through the English plural -s and second through reduplication. This type of double marking is called DOUBLING in Muysken (2000).

(82) *students-students* mampukah?
students-students afford?
'Can students afford it?'

However, doubling does happen more frequently in languages that are typologically distant as in Finnish-English code-switching where adpositions can be marked in both languages on a phrase (Muysken 2000, p. 104). In addition, it is a distinctive feature of ALT, and this type of switching is less frequent in our data.

In another example, the English noun *colour* was reduplicated to indicate plural (see 83).

(83) Tak payah *colour-colour*.
Not need colour-colour.
'No need for colours.'

The contrast in the direction of code-switching that we observed for nouns is also found for verbs and adjectives: there are about twice as many English verbs in Malay (57) than Malay verbs in English (33), and only two Malay adjectives in English, but 32 English adjectives in Malay. For adverbs, however, the distribution is less asymmetrical: 39 Malay adverbs in English and 21 English adverbs in Malay.

As for word order in the NP, English adjectives often appear in the canonical position before an English noun in a switched NP, as in (84).

(84) *Next week* kena tunjuk saya
Next week must show me
'Next week you must show me.'

English adjectives can also appear in predicative position, as in (85).

(85) Kenapa pula jadikan dia *greedy*?
Why moreover make him greedy?
'Why would you make him greedy?'

However, there are also cases where adjective placement appears to follow Malay rules, as in (86), where it is found before the Malay demonstrative *tu*. English adjectives can also be reduplicated, as in (87), and *punya* 'own' can appear between the English adjective and the English noun, as in (88).

(86) *Short* tu tak apa lagi.
Short that it's ok.
'As short as this is ok.'

(87) Tak payah *fancy-fancy*.
No need fancy-fancy.
'No need for fancy stuff.'

(88) Kenalah yang dah *raw* punya *analysis*.
It is necessary there is raw POSS analysis.
'It is necessary to make an analysis of your raw data.'

In summary, with respect to content words, it seems that speakers can draw freely on content words from both languages, although there are more English words in Malay utterances than vice versa. We did notice a preference for English words belonging to the semantic field of Linguistics, but further analyses would be required to establish whether there are specific semantic fields that are preferably expressed in English or Malay, as noticed by N. Abdullah (1975).

4.3.2. Closed Class Items

Pronouns

Most of the utterances with switches of pronouns consist of English pronouns in Malay sentences, as in (89), but there are a few cases where this is the other way around, as in (90).

(89) *I* tahu *you* tak baca lagi.
I know you not read more
'I know you don't read anymore.'

(90) Saya *just discuss then we can go straight to the next chapter*.
'I just discuss then we can go straight to the next chapter.'

Importantly, there are no instances of *I* being used as (in)direct objects, which was mentioned in Ozog (1987). It is possible that the occurrence of such forms is linked to the English language proficiency of the speakers, in that subject forms of pronouns for (in)direct objects only occur among speakers with lower proficiency levels. This was not the case for the two participants in our study. There was only one example (91), where *you*

is used as a possessive instead of standard English *your* (in combination with the possessive *punya*).

(91) Dia macam kalau you gaduh dengan *you* punya *girlfriend*
He like if you fight with you POSS girlfriend
'He's like, if you have a fight with your girlfriend.'

Demonstratives

There are also several instances of Malay demonstratives ((i)*tu* 'that' and (i)*ni* this) switched on their own in a completely English utterance, as in (92) and (93).

(92) *At least relate it to the study* tu.
At least relate it to the study that
'At least relate it to that study.'

(93) *What did you do in your discussion* ni?
What did you do in your discussion this
'What did you do in your discussion?'

These switches of demonstratives are similar to those found in Ozog (1987), who suggests they may function as discourse markers rather than as demonstratives. An alternative view would be to see this as a case of doubling, because in many cases the NP is accompanied by an English determiner which precedes the head and a Malay determiner which follows the head. Our examples are a little different from those of Ozog, because in (92), the speaker uses both the English definite article *the* and Malay *tu*. This combination is not mentioned in Ozog's overview. However, further research would be needed to determine the exact conditions under which English and Malay determiners can co-occur in one NP.

The Aspectual Marker Dah

Finally, the data contain many examples of the use of the aspectual marker *(su)dah* 'already', which is frequently combined with an English verb to mark the fact that an action has been completed, as in (94), where *dah* precedes the English verb *explain*.

(94) *Even I* dah *explain* pun *maybe reader* tak faham *so why not if I give*
Even I already explain even maybe reader not understand so why not if I give
example
example
'Even though I had already explained it, maybe the reader does not understand, so why don't I give (an) example' or 2.'

But *dah* appears to fulfill a wide range of functions, including adverbial ones, as in (95)

(95) *You* dah tahu
You already know
'You already know.'

It seems that its use in code-switching is similar to the uses described in Ozog (1987).

Modal Verbs

Our data contain many switches of Malay modal verbs. Here we will concentrate on the syntactic position of these modals, as an analysis of the subtle shades of meaning they express is beyond the scope of the current paper. The modals often occur in utterances in which all other words are English, as in (96), where the Malay modal verb *boleh* 'can' is the only Malay word in the utterance. In (97), by contrast, it is used in combination with a Malay lexical verb while the remaining words are English.

(96) *So that when you look at the answer key, you* boleh *guess how actually to get*
So that when you look at the answer key, you can guess how actually to get
that particular answer.
that particular answer.
'So that when you look at the answer key, you can guess how to actually get
that particular answer.'

(97) *These are the kind of* *things that you* boleh buat *research*-lah.
These are the kind of things that you can do research-DM
'These are the kind of things that you can do research on.'

Other Malay modals in English utterances include kena 'must', and mesti 'must'. Kena was sometimes used on its own, as in (98), sometimes in combination with an adverb, as in (99), while in (100) both *mesti* and *kena* are used in the same utterance. This combination of two modal expressions is not mentioned in other code-switching sources.

(98) *So one thing about this* *question* *is that your example* *you* kena *write* *the whole thing*.
So one thing about this question is that your example you must write the whole thing.
'So one thing about this question is that in your example you must write the whole thing.'

(99) *You* kena betul-betul *state what is your population*.
You must really state what is your population.
'You must really state what is the population.'

(100) *Portfolio* mesti-lah kena ada *cover page*.
Portfolio must-DM must there is cover page
'The portfolio must absolutely contain a cover page.'

Negation

As in other data sets, the negative marker *tak* (*tidak*) is used for negation in utterances which consist entirely or partly of English words, as in (101).

(101) tak *make sense*
not make sense
'does not make sense.'

Tak does not always appear in the same position. In (101) it appears before the compound verb *make sense*, which is the position in which it is found most frequently, but we also have an example where *tak* appears after the verb (see (102)).

(102) *Remember* tak *last time I mentioned to you that your article has to be . . .*
Remember not last time I mentioned to you that your article has to be . . .
what type of article
what type of article
'Do you not remember that I mentioned last time to you that your article had to
be . . . what type of article?'

Conjunctions and Discourse Markers

The English conjunction most frequently found in Malay utterances is *so*, as in (103).

(103) *So* kalau *you* baca *the next sentence*
So if you read the next sentence
'So if you read the next sentence.'

But and *because* are far less frequent in our data. In the opposite direction it is *tapi* 'but', *sebab* 'because', as in (100), and *kalau* 'if', as in (103), that are most frequent.

(104) Sebab *you are* *focusing* *on teaching* *them in school*
because you are focusing on teaching them in school
'Because you are focusing on teaching them in school.'

(105) Tapi *you have to contact me*-lah.
But you have to contact me-DM
'But you have to contact me.'

As for discourse markers, *lah* is used very frequently in our data, generally at the end of the utterance, and sometimes in combination with *kan*, as in (106), or with English discourse markers such as *of course*, as in (107).

(106) *And every each of these statement must* *be in the form of problem*-lah kan?
And every each of the statement must be in the form of problem- DM DM
'And each and every single of these statements must be in the form of a problem statement.'

(107) Tapi *of course* kena ada *data* dulu lah.
but of course must there be data first DM
'But of course there must be data first.'

Finally, *pun* 'even' is used in a variety of functions in code-switched utterances, as in (108).

(108) *So the article* yang *you submit* pun *cannot* *print a new one and submit*.
So the article that you submit even cannot print a new one and submit.
'So even when you have already submitted the article, you cannot print a new one and submit it.'

The Relative Clause Marker Yang

Another Malay function word that appears regularly in an English context is the relative clause marker *yang* 'who/that', which can be used to mark the subject or the object of a relative clause in Malay (Percillier 2016). In our data, *yang* can be used in a context that is entirely English, as in (109), where the head of the relative clause (*main idea*) is in English, and the subject of the relative clause (*you*) as well. *Yang* marks the object of the relative clause. As the relative clause which starts with *yang* in (109) also contains function words from Malay and English, namely the modal *kena* 'must', and the English pronoun *you*, function words come from both languages in this utterance, and the same is true for content words. Though most content words come from English, there is also one from Malay (*sendiri* 'alone'). The surface word order is the same in both languages.

(109) *Stated main idea* yang *you* kena *create* sendiri.
Stated main idea that you must create alone.
'Stated main idea that you must create on your own.'

While we did not find examples of switches of relative clause markers in the available academic literature on Malay-English code-switching, there are some in the global web-based English database (https://www.english-corpora.org/glowbe/ accessed on 20 October 2021). In some of these, *yang* is used as a relative pronoun on its own in an utterance which is completely English, as in (110), where it has an inanimate antecedent (*food*). In other contexts it is used at the start of a relative clause which is entirely in Malay, as in (111). Here the antecedent is animate (*the only one*). In both these examples it marks the subject but in (109) *yang* is used to mark the direct object of the relative clause, and its antecedent is *main idea*.

(110) *Out of all vendors, the food* yang *managed to capture my attention is* Mr Siew Bao.
Out of all vendors, the food which managed to capture my attention is Mr Siew Bao.
'Out of all vendors, the food which managed to capture my attention is that of Mr Siew Bao.' (Source: thebigsmallboy.com) [accessed on 21 October 2021]

(111) *And I am not the only one* yang rasa pelik
And I am not the only one who feels weird.
'And I am not the only one who feels weird.' (source: nannoor.blogspot.com) [accessed on 21 October 2021]

Further details about constructions involving relative clauses in Malaysian English, can be found in Percillier (2016).

Switching within a Word

There are no word-internal switches in the data.

In conclusion, we can say that apart from the Linguistics terminology, and the use of *contohnya* 'for example', which signals an upcoming explanation, the teachers' content word switches are very similar to those found in the available literature. This is also true for most switches of function words, except that no uses of the personal pronoun *I* as a direct object, indirect object or possessive were found in the teachers' switches. In addition, no word-internal switches were attested.

4.4. Do Switches of Malay Function Words in English Utterance Qualify as CLX?

The fourth research question focused on whether or not the examples of switches of function words can be interpreted as CLX. While Muysken (2014) includes switches of function words as one of the features that is indicative for CLX, there are several other reasons why this is indeed the correct analysis.

First of all, it is clear that in some utterances function words do not only come from one language (the ML) but can come from both languages, and the same is true for content words. Second, word order within a constituent can be partly or wholly shared between both languages, which is another key characteristic of CLX. In the case of modal verbs and the relative pronoun *yang*, the surface word order is completely shared between both languages. In addition, word order in a mixed NP can be flexible, in that switched adjectives appear sometimes before and sometimes after the noun. Thus, the grammars of both languages interact inside the NP, which is typical for CLX. While for demonstratives, modals and the relative pronoun *yang*, switching only takes place in one direction (Malay single word in English utterances), for pronouns switching can take place in both directions, even though switches of English pronouns in Malay are more frequent than switches in the other direction. The diversity of the switch patterns, as well as the fact that switching is bidirectional is another indication that the strategy used here is CLX. Note that INS is clearly unidirectional in many corpora, such as the Welsh-English Siarad corpus, where 98–99% of the monolingual and the bilingual clauses have Welsh as the ML and English as the embedded language (Deuchar 2020). In the data from the two teachers, there is also a clear asymmetry in switches of content words, in that there are more switches of English content words in Malay than switches in the opposite direction, but the proportion of switches of Malay words in English utterances is much higher than the proportion of Welsh words in English.

4.5. Explaining the Diversity in Patterns: Bilingual Optimization Strategies

Our final research question asked to what extent the patterns found could be explained on the basis of Muysken's bilingual optimization strategies. A good starting point could be the comments N. Abdullah (1975, pp. 31–32) made about the principles behind the code-switching patterns found in the data. She suggests they can be explained on the basis of the speakers' desire to economise or to abbreviate[11] or on the basis of 'economy of articulation', or as an avoidance strategy. In (112), for example, 'because she is frightened

of the cold' is represented in Malay with only two words *tak sejuk* 'not cold', or *disini* 'here'? in (113), which represents a full clause (*are you going to be here*) in English.

(112) *In winter she feels cold, in summer she feels cold*. Tapi *for shopping*, tak sejuk.
In winter she feels cold, in summer she feels cold. But for shopping, not cold.
'In winter she feels cold, in summer she feels cold. But for shopping it's never too cold.' (N. Abdullah 1975, p. 33)

(113) *How long* disini?
How long here?
'How long are you going to be here?' (N. Abdullah 1975, p. 31)

Conversely, according to the same author, the preference for some English expressions might also be related to the perceived efficiency of the English version by comparison with long-winded Malay translation equivalents (e.g., of *shopping* instead of *membeli-belah*).

Some evidence for N. Abdullah's (1975) 'brevity' argument can be found in Calude et al.'s (2020) study of the success of Maori loanwords in New Zealand English. They demonstrate that short Maori loanword (which were shorter than the English translation equivalent) were more likely to be adopted in New Zealand English than long Maori loanwords. Thus, for example, the relative shortness of the Maori word *iwi* by comparison with its translation equivalent 'tribe', made it a popular candidate for being adopted in New Zealand English.

The brevity or efficiency strategies mentioned by N. Abdullah and Calude et al. (2020) could be seen as examples of the universal principles behind language contact, mentioned in Muysken (2013). However, the use of Malay function words in English (and some English function words in Malay) is different. It is likely that speakers use Malay modal verbs such as *kena* 'must' in English utterances, because keeping two separate modality systems in memory is cognitively costly (see also Matras 2007 for a full discussion of the 'borrowability' of modality across languages). The speakers therefore opt to reduce the cognitive load by using CLX as an optimization strategy, which involves using Malay modal verbs in a shared grammatical frame. The multifunctionality of some Malay modal verbs, such as *kena*, constitutes an additional reason why they are attractive targets for importation into English. This importation could be facilitated by the lack of inflection on modal verbs in both languages. The outcome of language contact differs therefore from the scenario described by Moro (2015) who has argued that the Ambon Malay modal verbs have undergone semantic change as a result of contact with Dutch in the Netherlands, but does not mention borrowing or code-switching of Malay modal verbs in Dutch sentences.

The use of the relative pronoun *yang* in English sentences can be the result of an optimization strategy too, because *yang* fulfils all the functions of the range of relative pronouns *which, who(m)*, and *that*, which are very difficult for Malay learners of English (Wong and Chan 2005). Thus, the speaker does not need to keep in mind the different English forms and their grammatical functions when they opt for the Malay relative pronoun. A similar strategy (but in the opposite direction) might also explain bilinguals' preference for English pronouns of the first and second person, because the Malay translation equivalents are part of a very complex system with six different levels that encode respect, power and seniority (N. Abdullah 1975). For the third person pronoun, there are only two different options in Malay (a less formal variant *dia* and a more formal variant *beliau*), which makes this less complex than choosing among the pronouns for the first and second persons. In addition, there would be a significant cost to using the English pronouns of the third person because using these involves an obligatory choice between male (*he*, *him*) and female (*she*, *her*) pronouns, which is notoriously difficult for speakers whose first languages do not make that distinction (Dong et al. 2015). Keeping the Malay pronoun of the third person *dia* makes it possible to avoid these choices.

The use of Malay determiners/demonstratives in English utterances can also be seen as being motivated by bilingual optimization strategies, if Ozog (1987) is right in his analysis that these are in fact discourse markers. Inserting these into an English NP, makes

it possible for speakers to express the discourse-related functions that these markers fulfil in Malay. Some other switches (e.g., the use of bare nouns) can be the result of a strategy of neutrality (Muysken 1987), in that speakers choose switch points that least offend the structural or linear differences between the languages (in this case English articles).

Overall, we believe that the following key factors have contributed to the popularity of switches of function words in Malay-English code-switching. First of all, the basic word order of Malay and English is SVO (Cumming 1991), which means that the order in which the main constituents appear is similar in both languages. This makes it easier to find points at which switching between the languages can take place than in languages that have different basic word orders. Second, the two languages have been in contact for centuries, and convergence has taken place between both languages (Percillier 2016), which has further increased similarities. Third, the vocabularies overlap partly in that many English content words have been adopted into Malay (Azmi et al. 2016). When these are used they are likely to activate words from both languages, increasing the likelihood of further switches later in utterance. Fourth, there are few if any inflections in Malay, which facilitates switching of verbs (Muysken 2000), and this likely to also be true for modal verbs. It is possible that the absence of inflections also makes switching of nouns and adjectives in Malay easier. Fifth, there is no tradition of language separation, and the negative attitudes towards code-switching that are prevalent in Europe or North America, are unknown in Malaysia (David et al. 2009d).

Further research into the characteristics of individual switch types will be needed to provide additional explanations for the popularity of function words in Malay-English code-switching.

5. Conclusions

In this paper, we have seen that Malay-English code-switching is highly diverse. One of the most interesting aspects of switching in this language pair is that there is a wide range of switches of function words in Malay-English code-switching: these include modal verbs, determiners/demonstratives, personal pronouns and a relative pronoun, which are not frequently switched in other language pairs. In our data analysis we followed the methods proposed by Deuchar et al. (2007) to analyse our data. In the Malay-English data set, CLX was almost as frequent as INS, and for one of the two teachers it was the most frequent form of code-switching. A limitation of the current study was that the number of informants was small, and that their language proficiency in both languages was not measured. We cannot therefore confirm whether the frequency with which different types of code-switching occur in this data set is representative for the frequency with which these switches occur in the speech of other Malay-English bilinguals. However, we have found substantial evidence for the existence of CLX in this language pair, which was unexpected in view of the typological differences between English and Malay. In the contexts studied by Deuchar et al., and in Malaysia, depth of language contact may be a key variable explaining the frequency of CLX. In addition, in Malaysia, the lack of negative attitudes towards code-switching and the fact that there are very few (if any) inflections in Malay may have facilitated this type of code-switching. Finally, we have argued that the switches of function words are the result of bilingual optimization strategies aimed at matching L1 and L2 patterns where possible. This strategy is helpful in some contexts because it lightens the cognitive load of having to remember and use two different linguistic systems (Silva-Corvalán 1994). As the current paper sought to provide an overview of different patterns, we have not been able to treat all different patterns in great depth. We hope that further studies into the these fascinating switch patterns will also include experimental approaches, as these may enable more in-depth analyses of these phenomena.

Author Contributions: Conceptualization: S.M. and J.T.-D.; methodology: S.M.; software; formal analysis: S.M., J.T.-D. and Y.N.T.; writing—original draft preparation: S.M.; writing—review and editing: J.T.-D.; supervision: N.F. and J.T.-D. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Ethics Committee of the Institute of Education, University of Reading (22 October 2016).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Data can be obtained by contacting the two corresponding authors.

Acknowledgments: We would like to thank the two Malaysian teachers who took part in the experiment. We are also most grateful to Stefanie Pillai for sending us a copy of the paper by N. Abdullah (1975).

Conflicts of Interest: The authors declare no conflict of interest.

Notes

1 According to Prentice (1990), there are only two inflectional affixes in Malay, namely *meN-* and *di-*. See also Gil (2001) for a detailed treatment of prefixes in Malay and Indonesian dialects.

2 This construction is called a *Satzklammer* 'lit. sentence bracket' in traditional German grammar: the finite verb *haben* 'have' occupies the second position in the sentence and the non-finite verb (here the past participle *gemacht* 'made') occurs in sentence-final position. In Standard German, the PP *mit dem shopowner* 'with the shopowner' would appear within the Satzklammer, that is to the left of the past participle *gemacht*.

3 Ozog provides the gloss 'boy' but *seorang* is generally translated as 'person'. Ozog (1987) has glossed the example as a single NP, but an alternative analysis for this example is possible where the utterance is a complete clause, with *(i)tu*, the subject, and *form three seorang (pelajar 'student')*, the nominal predicate (see Hassan 2006, p. 19: for description of Malay predicates without verbal elements). In this interpretation, a more appropriate translation would be 'That is a form three boy'.

4 Sbb–abbreviation of *sebab* 'because'.

5 The meaning of *best* can vary according to the context according to Rasdi (2016). Here, the intended meaning is 'exciting' according to the author.

6 *mlm* is an abbreviation of malam 'night'

7 *lg* is an abbreviation of *lagi* 'again, more'

8 *mlm td* is an abbreviation of *malam tadi* 'last night'.

9 *inimlm* is an abbreviation of *ini malam* 'this night, tonight'.

10 The switch actually takes place just after the infinitival *to*, which is in fact part of the subordinate clause.

11 A similar argument was put forward by Turkish-German bilinguals on various occasions trying to explain Turkish-German code-switching patterns. In this language pair switching often took place between the German main clause and the Turkish subordinate clause, which consisted of constructions involving converbs or participles (see Treffers-Daller 2020 for details). The synthetic Turkish constructions were seen as more efficient than the analytical German constructions for subordination.

References

Abdullah, Imran H. 1993. *The Semantics of the Modal Auxiliaries of Malay*; Kuala Lumpur: Dewan Bahasa dan Pustaka.

Abdullah, Noor A. B. 1975. *Code-Switching among Bilingual Malays*. Unpublished Project. Kuala Lumpur: University of Malaya.

Ariffin, Kamisah, and Shameem Rafik-Galea. 2009. Code-switching as a communication device in conversation. *Language & Society Newsletter* 5: 1–19.

Ariffin, Kamisah. 2009. The Interaction of Language, Topic and Speakers: Code Switching in Classroom Discourse. In *Code-switching in Malaysia*. Edited by Maya. K. David, James McLellan, Shameem Rafik-Galea and Ain Nadzimah Abdullah. Frankfurt: Peter Lang, pp. 81–95.

Auer, Peter, Melissa Moyer, and Pieter Muysken. 2014. Connecting approaches to code-switching in West Africa. *International Journal of Bilingualism* 18: 447–54. [CrossRef]

Averintseva-Klisch, Maria. 2009. *Rechte Satzperipherie im Diskurs. Die NPRechtsversetzung im Deutschen*. Tübingen: Stauffenburg.

Azmi, Mohd Nazri Latif, Lidwina Teo Pik Ching, Norbahjah Binti Jamaludin, Muhammad Nur Haziq Bin Ramli, Muhammad Habibbullah Bin Razali, and Muhammad Ammar Yasser Bin Amram. 2016. The Comparisons and Contrasts Between English and Malay Languages. *English Review: Journal of English Education* 4: 209–18. [CrossRef]

Badiola, Lucia, Rodrigo Delgado, Ariane Sande, and Sara Stefanich. 2018. Code-switching attitudes and their effects on acceptability judgment tasks. *Linguistic Approaches to Bilingualism* 8: 5–24. [CrossRef]

Bolle, Jette. 1994. Sranan Tongo-Nederlands. Codewisseling en Ontlening. Master's thesis, University of Amsterdam, Amsterdam, The Netherlands.

Bukhari, Nurul I. B. A., Azu Farhana Anuar, Khairunnisa Mohad Khazin, and Tengku M. F. B. T. Abdul. 2015. English-Malay code-mixing innovation in Facebook among Malaysian University students. *Researchers World* 6: 1–10.
Calude, Andreea S., Steven Miller, and Mark Pagel. 2020. Modelling loanword success–a sociolinguistic quantitative study of Māori loanwords in New Zealand English. *Corpus Linguistics and Linguistic Theory* 16: 29–66. [CrossRef]
Clyne, Michael G. 2003. *Dynamics of Language Contact: English and Immigrant Languages*. Cambridge: Cambridge University Press.
Clyne, Michael G. 1967. *Transference and Triggering: Observations on the Language Assimilation of Postwar German-Speaking Migrants in Australia*. Leiden: Martinus Nijhoff.
Clyne, Michael G. 1987. Constraints on code-switching: How universal are they? *Linguistics* 25: 739–64. [CrossRef]
Cumming, Susanna. 1991. *Functional Change: The Case of Malay Constituent Order*. New York: Mouton de Gruyter.
David, Maya K., James McLellan, Shameen Rafik-Galea, and Ain Nadzimah Abdullah, eds. 2009a. *Code-Switching in Malaysia*. Frankfurt: Peter Lang.
David, Maya K., James McLellan, Kuang Ching Hei, and Ain Nadzimah Abdullah. 2009b. Code Switching Practices in Newspaper Headlines in Malaysia: A 50-year Comparative Study and Lexical Analysis. In *Code Switching in Malaysia*. Edited by Maya David, James McLellan, Shameen Rafik-Galea and Ain Nadzimah Abdullah. Frankfurt: Peter Lang, pp. 179–202.
David, Maya K., Kuang Ching Hei, James McLellan, and Hashim Fatimah. 2009c. Functions of Code Switching in the family domain in Malaysia. In *Code Switching in Malaysia*. Edited by Maya David, James McLellan, Shameen Rafik-Galea and Ain Nadzimah Abdullah. Frankfurt: Peter Lang, pp. 3–28.
David, Maya K., Wendy Yee Mei Tien, Ngeow Yeok Meng, and Gah Kah Hui. 2009d. Language choice and code switching of the elderly and the youth. *International Journal of the Sociology of Language* 200: 49–74. [CrossRef]
David, Maya. 2003. Role and functions of code-switching in Malaysian courtrooms. *Multilingua-Journal of Cross-Cultural and Interlanguage Communication* 22: 5–20. [CrossRef]
Dearden, Julie. 2014. *English as a Medium of Instruction—A Growing Global Phenomenon*. London: British Council. Available online: https://www.britishcouncil.org.br/sites/default/files/emi_a_growing_global_phenomenon.pdf (accessed on 1 November 2022).
Deuchar, Margaret, Pieter Muysken, and Sung-Lang Wang. 2007. Structured variation in codeswitching: Towards an empirically based typology of bilingual speech patterns. *International Journal of Bilingual Education and Bilingualism* 10: 298–340. [CrossRef]
Deuchar, Margaret. 2020. Code-switching in linguistics: A position paper. *Languages* 5: 22. [CrossRef]
Dewaele, Jean-Marc, and Wei Li. 2014. Intra-and inter-individual variation in self-reported code-switching patterns of adult multilinguals. *International Journal of Multilingualism* 11: 225–46. [CrossRef]
Dong, Yanping, Yun Wen, Xiaomeng Zeng, and Yifei Ji. 2015. Exploring the cause of English pronoun gender errors by Chinese learners of English: Evidence from the self-paced reading paradigm. *Journal of Psycholinguistic Research* 44: 733–47. [CrossRef] [PubMed]
Gil, David. 2001. The prefixes di-and N- in Malay/Indonesian dialects. In *The History and Typology of Western Austronesian Voice Systems*. Edited by Fay Wouk and Malcolm Ross. Canberra: The Australian National University, pp. 241–83.
Grosjean, François. 2001. The bilingual's language modes. In *One Mind, Two Languages: Bilingual Language Processing*. Edited by Janet L. Nicol. Oxford: Blackwell, pp. 1–22.
Habil, Hadina, and Shameem Rafik-Galea. 2009. Code-switching in organisational email discourse. In *Code Switching in Malaysia*. Edited by Maya David, James McLellan, Shameen Rafik-Galea and Ain Nadzimah Abdullah. Frankfurt: Peter Lang, pp. 169–78.
Hashim, Azirah. 2020. Malaysian English. In *The Handbook of Asian Englishes*. Edited by Kingsley Bolton, Werner Botha and Andy Kirkpatrick. Hoboken: John Wiley and Sons, pp. 373–97.
Hassan, Abdullah. 2006. *Morfologi*. Kuala Lumpur: PTS Professional Publishing.
Hofweber, Julia, Theo Marinis, and Jeanine Treffers-Daller. 2016. Effects of dense code-switching on executive control. *Linguistic Approaches to Bilingualism* 6: 648–668. [CrossRef]
Jacobson, Rodolfo. 2001. Aspects of scholarly language use in Malaysia: Switching codes in formal settings. *Contributions to The Sociology of Language* 84: 177–92.
Jaworska, Sylvia, and Christiana Themistocleous. 2018. Public discourses on multilingualism in the UK: Triangulating a corpus study with a sociolinguistic attitude survey. *Language in Society* 47: 57–88. [CrossRef]
Joshi, Aravind K. 1982. Processing of sentences with intra-sentential code-switching. Paper presented at the 9th Conference on Computational Linguistics, Prague, Czech Republic, July 5–10; vol. 1, pp. 145–50.
Karim, Nik S., Farid M. Onn, Hashim H. Musa, and Abdul H. Mahmood. 2008. *Tatabahasa Dewan Edisi Ketiga*; Kuala Lumpur: Dewan Bahasa dan Pustaka.
Kellerman, Eric. 1983. Now you see It, Now you don't. In *Language Transfer in Language Learning*. Edited by Susan M. Gass and Larry Selinker. Boston: Newbury House Publishers, pp. 112–34.
Klavans, Judith L. 1985. The syntax of code-switching: Spanish and English. In *Proceedings of the Linguistic Symposium on Romance Languages*. Amsterdam: Benjamins, vol. 14, pp. 213–31.
Koban, D. 2016. Attitudes towards oral code-switching among Turkish-English bilingual speakers in New York City. *Hacettepe Üniversitesi Türkiyat Araştırmaları Dergisi [Turkish Research Journal of Hacettepe University]* 24: 151–72.
Kroeger, Paul. 2014. Passive agents in Malay: The binding properties and discourse functions of agentive-nya. *NUSA: Linguistic Studies of Languages in and around Indonesia* 57: 5–29.
Labov, William. 1972. Some principles of linguistic methodology. *Language in Society* 1: 97–120. [CrossRef]

Lehtinen, Meri K. T. 1966. An Analysis of a Finnish-English Bilingual Corpus. Doctoral dissertation, Indiana University, Bloomington, IN, USA.
Levenshtein, Vladimir I. 1966. Binary codes capable of correcting deletions, insertions and reversals. *Soviet Physics Doklady* 10: 707–10.
MacWhinney, Brian. 2000. *The CHILDES Project: Tools for Analyzing Talk: Volume I: Transcription Format and Programs, Volume II: The Database*. New York: Psychology Press.
Majid, Sheikha. 2019. English to Malay and Back Again: An Analysis of Lecturers' Code Switching in English Classrooms. PhD dissertation, Institute of Education, University of Reading, Reading, UK.
Manan, Syed A., Maya K. David, Francisco P. Dumanig, and Khan Naqeebullah. 2015. Politics, economics and identity: Mapping the linguistic landscape of Kuala Lumpur, Malaysia. *International Journal of Multilingualism* 12: 31–50. [CrossRef]
Matras, Yaron. 2007. The borrowability of structural categories. In *Grammatical Borrowing in Cross-Linguistic Perspective*. Edited by Yaron Matras and Jeanette Sakel. Berlin: Mouton de Gruyter, pp. 31–74.
McClure, Erica F. 1977. *Aspects of Code-Switching in the Discourse of Bilingual Mexican-American Children*. Center for the Study of Reading Technical Report No. 044. Urbana-Champaigne: University of Illinois.
McLellan, James, and Rosalind Nojeg. 2009. Lexical choice and 3-way code switching in Bau-Jagoi Bidayuh discourse. In *Code Switching in Malaysia*. Edited by Maya K. David, James McLellan, Shameen Rafik-Galea and Ain Nadzimah Abdullah. Frankfurt: Peter Lang, pp. 49–61.
McLellan, James. 2009a. Keeping the Beast in Its Cage: Compartmentalization of English in the Malay World. In *Code Switching in Malaysia*. Edited by Maya K. David, James McLellan, Shameen Rafik-Galea and Ain Nadzimah Abdullah. Frankfurt: Peter Lang, pp. 203–24.
McLellan, James. 2009b. When two grammars coincide: Malay-English code-switching in public on-line discussion forum texts. *International Sociological Association RC25 Newsletter* 5: 1–22.
Moro, Francesca R. 2015. Modal categories in contact: The case of heritage Ambon Malay in the Netherlands. *Heritage Language Journal* 12: 271–91. [CrossRef]
Muysken, Pieter. 1987. Neutrality in code-mixing. In *Eigen en Vreemd. Handelingen van het 39e Nederlandse Filologencongres*. Amsterdam: VU Boekhandel/Uitgeverij, pp. 359–73. Available online: https://repository.ubn.ru.nl/bitstream/handle/2066/14621/3921.pdf (accessed 1 November 2022).
Muysken, Pieter. 2000. *Bilingual Speech: A Typology of Code-Mixing*. Cambridge: Cambridge University Press.
Muysken, Pieter. 2011. Bridges over troubled waters: Theoretical linguistics and multilingualism research. *Stellenbosch Papers in Linguistics* 40: 20–38. [CrossRef]
Muysken, Pieter. 2013. Language contact outcomes as the result of bilingual optimization strategies. *Bilingualism: Language and Cognition* 16: 709–30. [CrossRef]
Muysken, Pieter. 2014. Deja voodoo or new trails ahead. In *Linguistic Variation. Confronting Fact and Theory*. Edited by R. Torres Cacoullos, N. Dion and André Lapierre. London: Routledge, pp. 242–62.
Myers-Scotton, Carol. 1993. Common and uncommon ground: Social and structural factors in codeswitching. *Language in Society* 22: 475–503. [CrossRef]
Myers-Scotton, Carol. 1998. A way to dusty death: The matrix language turnover hypothesis. In *Endangered Languages: Language Loss and Community Response*. Edited by Lenore A. Grenoble and Lindsay J. Whaley. Cambridge: Cambridge University Press, pp. 289–316.
Myers-Scotton, Carol. 2002. *Contact Linguistics: Bilingual Encounters and Grammatical Outcomes*. Oxford: Oxford University Press.
Nadarajan, Shanthi. 2006. A crosslinguistic study of reduplication. *Journal of Second Language Acquisition and Teaching* 13: 39–53.
Nil, Zurina Mohamed, and Shamala Paramasivam. 2012. Code-Switching in Gol & Gincu. *Procedia-Social and Behavioral Sciences* 66: 169–75.
Othman, Normala. 2006. Current trends in pronoun usage among Malay speakers. Paper presented at Tenth Interntional Conference on Austranesian Linguistics, Palawan, Philippines, January 17–20.
Owens, Jonathan. 2005. Bare forms and lexical insertions in code-switching: A processing-based account. *Bilingualism: Language and Cognition* 8: 23–38. [CrossRef]
Ozog, A. Conrad K. 1987. The Syntax of the Mixed Language of Malay—English Bilinguals. *RELC Journal* 18: 72–90. [CrossRef]
Percillier, M. 2016. *World Englishes and Second Language Acquisition*. Amsterdam: John Benjamins Publishing.
Pfaff, Carol W. 1979. Constraints on language mixing: Intrasentential code-switching and borrowing in Spanish/English. *Language* 55: 291–318. [CrossRef]
Poplack, Shana, and James A. Walker. 2003. Review of Pieter Muysken, Bilingual speech: A typology of code-mixing. Cambridge: Cambridge University Press, 2000. *Journal of Linguistics* 39: 678–83. [CrossRef]
Poplack, Shana, and Marjory Meechan. 1995. Patterns of language mixture: Nominal structure in Wolof-French and Fongbe-French bilingual discourse. In *One Speaker, Two Languages: Cross-Disciplinary Perspectives on Code-Switching*. Edited by Lesley Milroy and Pieter Muysken. Cambridge: Cambridge University Press, pp. 199–231.
Poplack, Shana. 1980. Sometimes I'll start a sentence in spanish y termino en espanol: Toward a typology of code-switching. *Linguistics* 18: 581–618. [CrossRef]
Poplack, Shana. 2018. *Borrowing. Loanwords in the Speech Community and in the Grammar*. New York: Oxford University Press.

Prentice, David John. 1990. Malay (Indonesian and Malaysian). In *The Major Languages of East and South-East Asia*. Edited by Bernard Comrie. London: Routledge.
Rasdi, Nur I. B. 2016. Intrasentential Code-Switching among Malays in Malaysia on Facebook. Doctoral dissertation, University of Essex, Colchester, UK.
Sebba, Mark. 1998. A congruence approach to the syntax of codeswitching. *International Journal of Bilingualism* 2: 1–19. [CrossRef]
Silva-Corvalán, Carmen. 1994. *Language Contact and Change: Spanish in Los Angeles*. New York: Oxford University Press.
Smith, Mike Sharwood, and Eric Kellerman. 1986. Crosslinguistic influence in second language acquisition: An introduction. In *Crosslinguistic Influence in Second Language Acquisition*. New York: Pergamon.
Sneddon, James N. 2007. *Indonesian Reference Grammar*. Crows Nest: Allen & Unwin.
Talmy, Leonard. 2000. *Toward a Cognitive Semantics. Vol. I: Concept Structuring Systems*. Cambridge and London: MIT Press.
Tay, Li C., Mei Y. Chan, Ngee Thai Yap, and Bee E. Wong. 2016. Discourse particles in Malaysian English: What do they mean? *Bijdragen tot de taal-, land-en volkenkunde/Journal of the Humanities and Social Sciences of Southeast Asia* 172: 479–509. [CrossRef]
Treffers-Daller, Jeanine. 1994. *Mixing Two Languages: French-Dutch Contact in a Comparative Perspective*. Berlin: Walter de Gruyter.
Treffers-Daller, Jeanine. 2005. Evidence for insertional codemixing: Mixed compounds and French nominal groups in Brussels Dutch. *International Journal of Bilingualism* 9: 477–506. [CrossRef]
Treffers-Daller, Jeanine. 2009. Code-switching and transfer: An exploration of similarities and differences. In *The Cambridge Handbook of Linguistic Codeswitching*. Edited by Barbara E. Bullock and Jacqueline A. Toribio. Cambridge: Cambridge University Press, pp. 58–74.
Treffers-Daller, Jeanine. 2020. Turkish-German code-switching patterns revisited: What naturalistic data can(not) tell us. In *Advances in Contact Linguistics. In honour of Pieter Muysken*. Edited by Tonjes Norval Smith and Enoch O. Aboh. Amsterdam and Philadelphia: John Benjamins Publishing Company, pp. 237–60.
Treffers-Daller, Jeanine. Forthcoming. The Simple View of Borrowing and Code-switching. *The International Journal of Bilingualism*. *under view*.
Vollmann, Ralf, and Tek Wooi Soon. 2020. Language change and convergence in multilingual Malaysian Chinese. *Global Chinese* 6: 49–67. [CrossRef]
Wong, Angelina L. W. 2012. Code-Switching in Blogs of Older and Younger Malaysian Chinese. Master thesis, University of Malaya, Kuala Lumpur, Malaya.
Wong, Eng, and Heng Chan. 2005. English relative clauses: What Malay learners know and use. *Pertanika Journal of Social Science and Human* 13: 107–15.
Wong, Irene F. H. 1981. English in Malaysia. In *English for Cross-Cultural Communication*. Edited by Larry Smith. London: Palgrave Macmillan, pp. 94–107.